THIRD AUTUMN COURSE ON MATHEMATICAL ECOLOGY

(29 October - 16 November 1990)

"Towards a unified theory of pair formation"

C. CASTILLO-CHAVEZ
Cornell University
Centre for Applied Mathematics
Biometrics Unit
Ithaca, NY 14853-7801
U.S.A.

Towards a unified theory of pair formation

Stephen P. Blythe[1]	Carlos Castillo-Chavez[2]	Jeffrey S. Palmer	Mingyan Cheng
Dept. of Stat. &	332 Warren Hall	322 Warren Hall	337 Warren Hall
Modelling Science	Biometrics Unit/Center	Biometrics Unit	Biometrics Unit
University of Strathclyde	for Appl. Math./ Pop. &	Cornell University	Cornell University
Glasgow, Scotland	Development Program	Ithaca, NY 14853	Ithaca, NY 14853
Great Britain	Cornell University		
	Ithaca, NY 14853-7801		

Abstract

Sexually-transmitted diseases, such as gonorrhea, syphilis, herpes, and AIDS, are driven and maintained in populations, by epidemiological and sociological factors that are not completely understood. Despite the fact that the processes of pair formation (or social mixing) and dissolution play a crucial role in disease dynamics their incorporation into epidemiological models is quite recent. In this paper, we present a unified approach to pair formation for a population with an arbitrary number of "sexes." A new derivation is provided of the mixing formula of Busenberg and Castillo-Chavez and special cases, such as two- and one-sex models, are discussed. We illustrate how some of the mixing formulas that have appeared in the literature fit into our framework. We outline the results of some stochastic simulations and compare the averages of our simplest stochastic simulations to special deterministic solutions – proportionate mixing and Ross♥ solutions and provide an illustration of our approach in the context of genetic mating systems. Finally, we end with a discussion of female (or male) choice models.

Section 1. Introduction

Recent advances in modeling the epidemiology of sexually-transmitted diseases, and in particular the human immunodeficiency virus (HIV, the causative agent of AIDS), have produced some striking new results in the mathematical description of sexual mixing processes. In addition to the standard random mixing with heterogeneity in numbers of partners (called "proportionate mixing," *e.g.* Barbour 1978, Nold 1980, Hethcote and Yorke 1984, Anderson and May 1984, Dietz and Schenzle 1985, Castillo-Chavez *et al.* 1988, 1989), a version of the assortative mating structure familiar to population geneticists has been used (called "preferred mixing" in the STD literature; *e.g.* Nold 1980, Hethcote

---

[1] Current address: 322 Warren Hall; Biometrics Unit; Cornell University; Ithaca, NY 14853-7801

[2] To whom correspondance and information about reprints should be addressed

models *also*

and Yorke 1984, Jacquez et.al. 1988), and ~~models~~ with adaptive sexual behavior have ~~also~~ been suggested by Anderson *et al.* (1989). A number of particular-case "mixing functions" have also been proposed (e.g. Castillo-Chavez and Blythe 1989; Hyman and Stanley 1989; Koopman et.al. 1989), but the field has been significantly opened up by Busenberg and Castillo-Chavez (1989, 1990) who generalized the specific case of Blythe and Castillo-Chavez (1989)--called "like-with-like mixing"--and obtained a representation theorem stating that all mixing functions may be expressed in a special form (Blythe and Castillo-Chavez (1990) show how many of the particular cases fit into the general form). Specific two-sex, age-structured, multi-partner results have also been obtained (Castillo-Chavez and Busenberg, 1990) as well as solutions for arbitrarily connected groups of any type (Blythe, 1990).

In its general form, the results of Busenberg and Castillo-Chavez (1989, 1990) deal with sexual contacts (*i.e.* partners) per unit time in a population comprised of N groups, in the ith of which there are $T_i(t)$ individuals with average number of partners $c_i(t)$ at time t. A valid description of the mixing process is an N × N matrix of probabilities $p(t)$ where $p_{ij}(t)$ is the probability that an individual in group i has a partner in group j, at time t. There are four fundamental constraints on $p(t)$ which specify solutions:

(i) $\quad 0 \leq p_{ij}(t) \leq 1 \quad$, all i, j and t

(ii) $\quad \sum_{j=1}^{N} p_{ij}(t) = 1 \quad$, all i and t

(iii) $\quad c_i(t)T_i(t)p_{ij}(t) = c_j(t)T_j(t)p_{ji}(t) \quad$, all i, j and t

(iv) $\quad c_i(t)T_i(t)c_j(t)T_j(t) <=> p_{ij}(t) = p_{ji}(t) = 0 \quad$, all i and j, and any t.

Constraints (i) and (ii) simply make **p** a stochastic matrix, (iii) enforces conservation of the number of new pairings per unit time between individuals in groups, and (iv) says that individuals in momentarily empty or inactive groups cannot have partners. The representation theorem states that *any* **p**, a solution to the problem specified by constraints (i) - (iv), may be written in the form:

$$p_{ij}(t) = \bar{p}_j(t) \left[\frac{R_i(t)R_j(t)}{V(t)} + \phi_{ij} \right] \quad \text{, all i, j, t} \tag{1}$$

where

$$\bar{p}_j(t) = \frac{c_j(t)T_j(t)}{\sum_{k=1}^{N} c_k(t)T_k(t)} \quad \text{, all j} \tag{2}$$

represents random or proportionate mixing between groups,

$$R_i(t) = 1 - \sum_{k=1}^{N} \bar{p}_k(t)\phi_{ik} \quad , \text{all } i \tag{3}$$

and

$$V(t) = \sum_{k=1}^{N} \bar{p}_k(t)R_k(t) \quad . \tag{4}$$

Here, ϕ is an N × N matrix with

$$\phi_{ij} = \phi_{ji} \tag{5}$$

that provides *a measure of mutual preference, or affinity* for sexual partners between pairs of groups (see Blythe and Castillo-Chavez 1989, Castillo-Chavez and Blythe 1989, and Castillo-Chavez *et al.* 1990a). A simple new diagramatic derivation of formula (1) is provided in Section 2, while the proof that all possible mixing solutions (for age- and socially-structured populations) can be written in this fashion may be found in Busenberg and Castillo-Chavez (1989, 1990). In Section 5, we discussed several particular mixing solutions, in the context of this framework, that are found in the literature.

This framework has been incorporated into models for the sexual transmission of HIV among homosexually-active populations (see Castillo-Chavez *et al.* 1989a, b) and has already produced new theoretical results (see Castillo-Chavez *et al.* 1989a, Huang 1989, and Huang *et al.* 1990). Our recent work has concentrated on the development of statistical methods for the estimation of the matrix ϕ (Blythe, Castillo-Chavez and Casella 1990). This framework has also been incorporated into demographic models that follow pairs of individuals (see Castillo-Chavez *et al.* 1990a) and we discussed models of this type in Section 3. To illustrate the flexibility of this approach we introduce our framework in the context of a "classical" epidemic model, for the homosexual transmission of gonorrhea among N-interacting populations, using as few definitions and equations as possible.

Let $S_i(t)$ and $I_i(t)$ respectively denote the number of susceptibles and infecteds in the ith group, at time t. Let Λ_i denote the rate of influx (recruitment) of new susceptibles to the ith group, and let $1/\mu$ and $1/\sigma$ be the average duration of a sexual "lifetime" and the average duration of the infected phase, respectively. If $B_i(t)$ is the incidence rate (of infections) in the ith group, then we may write

$$\frac{dS_i(t)}{dt} = \Lambda_i - B_i(t) - \mu S_i(t) + \sigma I_i(t) \tag{6}$$

$$\frac{dI_i(t)}{dt} = B_i(t) - (\mu+\sigma) I_i(t) \tag{7}$$

for i = 1, 2, ..., N. We of course require initial conditions $S_i(0) > 0$, $I_i(0) \geq 0$ for all i. The incidence

rates are given by

$$B_i(t) = S_i(t) \sum_{j=1}^{N} \beta \, p_{ij}(t) c_i(t) \frac{I_j(t)}{T_j(t)} , \tag{8}$$

which is interpreted as follows. Each individual in group i has $c_i(t)$ partners per unit time at time t. Of these, a fraction $p_{ij}(t)$ come from group j (j = 1, 2, ..., N), and of these, a fraction $I_j(t)/T_j(t)$ are infected at time t ($T_j(t) = S_j(t) + I_j(t)$, ie., the total population of the ith group). In this simple illustrative example, there is assumed to be a constant probability β of a susceptible person becoming infected during a partnership with an infected person. Thus the summation term on the RHS of Equation (8) is the probability per unit time of a susceptible person in group i becoming infected at time t, and hence $B_i(t)$ is the total rate of new infections occurring in group i at time t. In this simplified form, if we prescribe the p matrix, then we have a complete specification of the STD epidemic model. Fig (1) shows how the p(t) surface evolves with time. The surfaces are derived from Eq (1)-(8) with arbitrary parameter values, and a special case of mixing parameters ($\phi_{ii} = a$, and $\phi_{ij} = b$ if $i \neq j$). Models of this type, but modified to apply to the sexual transmission of AIDS (*e.g.* with disease-induced mortality, and no recovery term) have been analyzed by Castillo-Chavez *et al.* (1989a, b), Huang (1989), and Huang *et al.* (1990). Classical epidemiological models that incorporate age-structure and a generalized mixing framework based on the same axiomatic system have been developed and partially analyzed by Busenberg and Castillo-Chavez (1989, 1990). In Busenberg and Castillo-Chavez (1990) a table providing nine simple explicit forms of age-dependent mixing functions is also provided. We do not incorporate examples of this type because our objective is only to present a unified approach to mixing in its simpler form.

The study of biological questions has dictated our approaches and not the other way around (as is still too common in mathematics). Researchers that favor stochastic approaches may question their absence in addressing these biological issues. We have *not* neglected stochastic approaches, and are currently looking at stochastic simulations of the processes of pair-formation and dissolution, and at stochastic differential equation infection models, in the context of sexually transmitted diseases. Here (Section 6), we describe our efforts to address these issues using probabilistic approaches and their relationship to our deterministic framework.

Section 2. A Diagramatic View of General Mixing in a Single-Sex Population

We offer a new biological derivation for the General Mixing solution (1) in a homosexually active population. This idea is described for N interacting subgroups and is illustrated for the case N=2 in

Figure 2. This derivation is also valid in more general contexts see Castillo-Chavez *et al.* (1990b).

Examine a typical individual from subgroup i (with partnership acquisition rate $c_i(t)$ per unit time). The General Mixing formula (1) together with axiom (i) imply that $p_{ij}(t) \geq \bar{p}_j(t)\phi_{ij}(t)$. Thus, an average individual in subgroup i forms partnerships with individuals from subgroup j at a minimum rate of

$$c_i\phi_{ij}(t)\bar{p}_j(t). \tag{9}$$

Consequently, summing over all subgroups, we conclude that a typical member of this population has formed within-group partnerships at a minimal rate of

$$c_i\sum_{k=1}^{N}\phi_{ik}(t)\bar{p}_k(t) = c_i[1 - R_i(t)]. \tag{10}$$

The remaining rates $c_iR_i(t)$ (i = 1, 2,..., N) will be distributed using a different mechanism.

Consider a population divided into N distinct subgroups, each of size $T_j(t)$, but which forms partnerships at the (reduced) rate $c_jR_j(t)$. Assume Proportionate Mixing among the subgroups in this "new" population (same population but with reduced rates for pairing). Then an average individual from subgroup i (mixing at random) forms partnerships with individuals from subgroup j at the rate of

$$c_iR_i(t)\frac{c_jR_j(t)T_j(t)}{\displaystyle\sum_{k=1}^{N}c_k\bar{p}_k(t)R_k(t)} = \frac{c_iR_i(t)R_j(t)\bar{p}_j(t)}{\displaystyle\sum_{k=1}^{N}\bar{p}_k(t)R_k(t)} \tag{11}$$

per unit time. Adding together the results in (9) and (11) above, we see that a typical individual from subgroup i has formed partnerships with members of group j at the rate

$$c_i\left[\phi_{ij}(t)\bar{p}_j(t) + \frac{R_i(t)R_j(t)\bar{p}_j(t)}{\displaystyle\sum_{k=1}^{N}\bar{p}_k(t)R_k(t)}\right] = c_i\bar{p}_j(t)\left[\frac{R_i(t)R_j(t)}{\displaystyle\sum_{k=1}^{N}\bar{p}_k(t)R_k(t)} + \phi_{ij}(t)\right] = c_iP_{ij}(t). \tag{12}$$

Note that in the above we have made no assumption about the $\{\phi_{ij}\}$. They could in fact be density or frequency dependent, or even *explicitly* time dependent. Constraints on the range of ϕ's that we use are provided that the following condition is satisfied (see Busenberg and Castillo-Chavez 1989, 1990): $R_i \geq 0$ (i = 1, 2,..., N) with at least one of the R_i's > 0.

5

Since all mixing functions can be written in the form of General Mixing, this provides a simple biological explanation of the way in which mixing between "subgroups" of a single-sex population can be described. However, this does *not* imply that mixing occurs in this way among individuals, a problem for which pair-formation models are more appropriate. The biology of General Mixing may be thought of as reserving a fraction, $\phi_{ij}(t)\bar{p}_j(t)$, of subgroup i's rate of pair formation for interactions with subgroup j. This fraction is the product of $\phi_{ij}(t)$, a measure of preference or affinity, and $\bar{p}_j(t)$, a measure of relative availability of partners. The remaining rate of pair formation is distributed at random (proportionate mixing) among the various subgroups, which are now mixing at a reduced rate $c_i R_i(t)$. We finally note that, in general, nothing prevents us from assuming that the rate in (9) is negative. This assumption can also provide us with an alternative representation theorem where negative values for the ϕ's are allowed (some new restrictions on the R_i's are required). These representations are further explored in Castillo-Chavez *et al.* (1990b).

Section 3. Pair formation

All of the mixing solutions are written in terms of contact distributions - individuals in each of the groups comprising the population all take the same number of partners (contacts) per unit of time, and this number varies between groups, and possibly with time. An alternative approach, using models that follow the dynamics of pair fromation and dissolution, has been studied by Kendall (1949), Keyfitz (1949), Parlett (1972), Fredrickson (1971), McFarland (1972), Pollard (1973), Gimelfarb (1988a, b), Dietz and Hadeler (1988), Dietz (1988), Hadeler (1989a, b, 1990), Waldstätter (1989), Castillo-Chavez (1989), Busenberg and Castillo-Chavez (1989, 1990), Castillo-Chavez and Busenberg (1990), and Castillo-Chavez *et al.* (1990a, b), Blythe and Castillo-Chavez (1990), and Blythe (1990), with heterogeneity arising because different individuals have different rates for the processes of pair formation and dissolution.

The mixing solution due to Castillo-Chavez and Busenberg has been shown to be the general solution for the one-sex mixing problem (Busenberg and Castillo-Chavez, 1990), and the related form (this section) has been derived for, and shown to be the general solution of the two-sex mixing problem (Castillo-Chavez and Busenberg, 1989, 1990). This solution provides a unified approach to the problem of pair-formation or social mixing as it can be incorporated into classical demographic and epidemiological models as well as into those that follow pairs (see Castillo-Chavez and Busenberg 1990, and Castillo-Chavez *et al.* 1990). In these articles it is shown that if we fix the rate of pair-formation and let the rate of pair-dissolution go to infinity then the pair-formation models become the classical mixing models. Consequently, useful comparisons between stochastic and deterministic mixing models,

for large populations, are handled more appropriately in the context of models that follow pairs (see Section 6). We observe further, that our pair-formation formalism may be applied in the study of genetic systems, and in Section 7 we provide some preliminary results using a very simple population genetics model.

To provide a context for the sexual interactions of a heterosexual population, we introduce a two-sex model with variable population size for the transmission dynamics of gonorrhea that follows pairs of individuals. Traditional gonorrhea models (see Hethcote and Yorke, 1984) have assumed that the mixing subpopulations have constant size. This assumption may be very useful when we deal with the relative evaluation of control strategies (*loc. cit.*). However, this assumption is not appropriate in situations in which we wish to evaluate the impact of different mixing patterns in disease dynamics. The assumption of interacting populations of constant size leads to time-independent mixing probabilities (i.e. constant contact matrices) and hence to mixing patterns that are valid only for populations that have already reached a steady state.

We consider a population of heterosexually active individuals. This population is divided into classes or subpopulations. Classes can be identified by sex, race, socio-economic background, average degree of sexual activity, etc. Models that incorporate factors such as chronological age, age of infection, variable infectivity, and partnership duration can be found in our earlier work (see Blythe and Anderson 1989; Busenberg and Castillo-Chavez, 1989, 1990; Thieme and Castillo-Chavez 1989, 1990). Fig (3) is a schematic of the main features of the model. We consider N sexually active populations of females and L sexually active populations of males. Each population is divided into two epidemiological classes for single individuals: $f_j(t)$ and $m_i(t)$ (single susceptible females and males, i.e., uninfected and sexually-active, at time t); $F_j(t)$ and $M_i(t)$ (single infected females and males, at time t); for j = 1,...,N and i = 1,...,L. Hence the sexually active single individuals of each sex and each subpopulation at time t are represented by $T_j^f(t) = f_j(t) + F_j(t)$ and $T_i^m(t) = m_i(t) + M_i(t)$. The epidemiological classes for pairs are given by $P_{ji}^{fm}(t)$, $P_{ji}^{Fm}(t)$), $P_{ji}^{fM}(t)$, $P_{ji}^{FM}(t)$, and naturally, transmission can only occur among those individuals in pairs $P_{ji}^{Fm}(t)$) or $P_{ji}^{fM}(t)$. We note that $P_{ji}^{fM}(t) = P_{ij}^{Mf}(t)$ and consequently we need only to consider four types of pairs. We assume that the transmission probability per unit time is constant within each pair containing one infected individual, and denote by δ_M and δ_F the rates for male-to-female and female-to-male transmission, respectively. We further assume that the percapita recovery rates are given by the constants γ_M and γ_F, and that these rates are independent of whether or not the individuals are paired. The per capita dissolution rates are given by the constants σ_{fm}, σ_{fM}, σ_{Fm}, and σ_{FM}, and the per capita removal rates from sexual activity (due to death or other causes not previously considered) are given by μ_f, μ_m,

μ_F, and μ_M.

To simplify the renewal processes, we assume that $\Lambda_j{}^f$ and $\Lambda_i{}^m$ denote the "recruitment" rates (assumed constant) of single (assumed uninfected) individuals.

Of course, this model is not fully described until we provide explicit expressions for the interactions terms, that is a set of mixing or pair-formation probabilities $\{p_{ij}(t)$ and $q_{ji}(t)$: $i = 1,...,L$ and $j = 1,...,N\}$. So we let

$p_{ij}(t)$: denotes the proportion of partnerships of males in group i with females in group j at time t, and

$q_{ji}(t)$: denotes the proportion of partnerships of females in group j with males in group i at time t,

and if

c_i : denotes the average (constant) rate at which males in group i form partnerships with females in any group. It is the ith-group rate of (male) pair-formation, and

b_j : denotes the average (constant) rate at which females in group j form partnerships with males in any group. It is the jth-group rate of (female) pair-formation.

then these proportions satisfy the axioms stated in the following definition.

<u>Def</u> $(p_{ij}(t), q_{ji}(t))$ is called a mixing/pair-formation matrix iff it satisfies the following properties (at all times):

(I) $0 \le p_{ij} \le 1,$ $0 \le q_{ji} \le 1,$

(II) $\sum\limits_{j=1}^{N} p_{ij} = 1 = \sum\limits_{i=1}^{L} q_{ji},$

(III) $c_i T_i{}^m p_{ij} = b_j T_j{}^f q_{ji}, \qquad i = 1, \cdots, L, \quad j = 1, \cdots, N.$

(IV)   If for some i, $0 \le i \le L$ and/or some j, $0 \le j \le N$ we have that $c_i b_j T_i{}^m T_j{}^f = 0$, then we define $p_{ij} \equiv q_{ji} \equiv 0$.

Note that (III) can be viewed as a conservation of partnerships law or a group reversibility property, while (IV) asserts that the mixing of "non-existing" or non-sexually active subpopulations cannot be arbitrarily defined. For the gonorrhea model, and most deterministic models for STD's, subpopulations that are sexually active do not become extinct and remain sexually active for all time. We introduce the following notation for the "restricted" mixing functions:

$$q_{ji}{}^{xm}(t) = \frac{m_i}{M_i + m_i} \; q_{ji}(t), \tag{13}$$

$$q_{ji}{}^{xM}(t) = \frac{M_i}{M_i + m_i} \; q_{ji}(t), \tag{14}$$

$$p_{ij}{}^{yf}(t) = \frac{f_i}{F_i + f_i} \; p_{ij}(t), \tag{15}$$

$$p_{ij}{}^{yF}(t) = \frac{F_i}{F_i + f_i} \; p_{ij}(t), \tag{16}$$

where x = f or F and y = m or M.

With these definitions we can, with the aid of Figure 3, write the following pair-formation model for the transmission dynamics of gonorrhea:

$$\frac{df_j(t)}{dt} = \Lambda_j{}^f - [b_j + \mu_f]f_j(t) + \gamma_F F_j(t) +$$

$$[\mu_M + \sigma_{fM}] \sum_{i=1}^{L} P_{ji}{}^{fM}(t) + [\mu_m + \sigma_{fm}] \sum_{i=1}^{L} P_{ji}{}^{fm}(t), \tag{17}$$

$$\frac{dF_j(t)}{dt} = -[b_j + \gamma_f + \mu_F] \; F_j(t) - [\mu_m + \sigma_{Fm}] \sum_{i=1}^{L} P_{ji}{}^{Fm}(t) + [\mu_M + \sigma_{FM}] \sum_{i=1}^{L} P_{ji}{}^{FM}(t), \tag{18}$$

$$\frac{dm_i(t)}{dt} = \Lambda_i{}^m - [c_i + \mu_m] \; m_i(t) + \gamma_M M_i(t) +$$

$$[\mu_F + \sigma_{Fm}] \sum_{j=1}^{N} P_{ji}{}^{Fm}(t) + [\mu_f + \sigma_{fm}] \sum_{j=1}^{N} P_{ji}{}^{fm}(t), \tag{19}$$

$$\frac{dM_i(t)}{dt} = -[c_i + \gamma_M + \mu_M] \; M_i(t) + [\mu_f + \sigma_{fM}] \sum_{j=1}^{N} P_{ji}{}^{fM}(t) + [\mu_F + \sigma_{FM}] \sum_{j=1}^{N} P_{ji}{}^{FM}(t) \tag{20}$$

$$\frac{dP_{ji}{}^{fm}(t)}{dt} = b_j \cdot [q_{ji}{}^{fm}(t)] \; f_j(t) - [\mu_f + \mu_m + \sigma_{mf}]P_{ji}{}^{fm}(t) +$$

$$[\gamma_M] \; P_{fM} + [\gamma_F] \; P_{Fm} + [\gamma_F + \gamma_M + \gamma_F \gamma_M]P_{FM}(t) \; , \tag{21}$$

$$\frac{dP_{ji}{}^{Fm}(t)}{dt} = b_j\,[q_{ji}{}^{Fm}(t)]\,F_j(t) - [\mu_F + \mu_m + \sigma_{mF} + \delta_M + \gamma_F]P_{ji}{}^{Fm}(t) +$$

$$[\gamma_M]P_{ji}{}^{FM}(t) - [\gamma_F]P_{ji}{}^{Fm}(t) \tag{22}$$

$$\frac{dP_{ji}{}^{fM}(t)}{dt} = b_j[q_{ji}{}^{fM}(t)]\,f_j(t) - [\mu_f + \mu_M + \sigma_{fM} + \delta_F + \delta_M]P_{ji}{}^{fM}(t) +$$

$$[\gamma_F]P_{ji}{}^{FM}(t) - [\gamma_M]P_{ji}{}^{Fm}(t) \tag{23}$$

$$\frac{dP_{ji}{}^{FM}(t)}{dt} = b_j[q_{ji}{}^{FM}(t)]\,F_j(t) - [\mu_F + \mu_M + \sigma_{MF} + \gamma_M + \gamma_F + \gamma_F\gamma_M]P_{ji}{}^{FM}(t)$$

$$+[\gamma_F]P_{ji}{}^{Fm}(t) + [\gamma_M]P_{ji}{}^{fM}(t), \tag{24}$$

$i = 1,...,L$ and $j = 1,...,N$.

Remark

We note that axiom (III) has to be satisfied for all times and that this includes $t = 0$. Therefore, the initial conditions cannot be arbitrary. In the past, very little attention has been paid to this constraint. We will use this constraint when we discuss, in Section 8, our female (male) choice models.

Before, we proceed some insight is needed as to the explicit nature of the mixing probabilities. To this end, we compute the Ross solutions:

<u>Def</u>   A two-sex mixing/pair-formation function is called separable iff

$$P_{ij} = P_i P_j \quad and \quad q_{ji} = Q_j Q_i \, ,$$

for some appropriate functions P_i and Q_i. This definition lead us to the following useful characterization of two-sex separable mixing function.

<u>Theorem 1:</u>   The only separable solutions are given by the set of Ross solutions $\{(\bar{p}_j, \bar{q}_i)\}$ where

$$\bar{p}_j = \frac{b_j T_j{}^f}{\sum\limits_{i=1}^{L} c_i T_i{}^m}, \qquad \bar{q}_i = \frac{c_i T_i{}^m}{\sum\limits_{j=1}^{N} b_j T_j{}^f}; \qquad j = 1, \cdots, N \quad and \quad i = 1, \cdots, L \,.$$

Ross solutions can be used to characterize all possible mixing solutions. The following result was established by Castillo-Chavez and Busenberg (1990):

Theorem 2. Let $\{\phi_{ij}^{m}\}$ and $\{\phi_{ji}^{f}\}$ be two nonnegative matrices. Let $\ell_{i}^{m} \equiv \sum\limits_{k=1}^{N} \bar{p}_{k}\phi_{ik}^{m}$ and $\ell_{j}^{f} \equiv \sum\limits_{k=1}^{L} \bar{q}_{k}\phi_{jk}^{f}$ where $\{(\bar{p}_{j},\bar{q}_{i})\ j = 1,...,N$ and $i = 1,...,L\ \}$ denotes the set composed of Ross's solutions. We also let $R_{i}^{m} \equiv 1 - \ell_{i}^{m}$, $i = 1, \cdots, L$ and $R_{j}^{f} \equiv 1 - \ell_{j}^{f}$, $j = 1, \cdots, N$, and assume that ϕ_{ij}^{m} and ϕ_{ji}^{f} are chosen in such a way that R_{i}^{m} and R_{j}^{f} remain nonnegative for all time. We further assume that

$$\sum_{i=1}^{L} \ell_{i}^{m}\bar{p}_{i} = \sum_{i=1}^{L}\sum_{k=1}^{N} \bar{p}_{k}\phi_{ik}^{m}\bar{p}_{i} < 1,$$

and

$$\sum_{j=1}^{N} \ell_{j}^{f}\bar{q}_{j} = \sum_{j=1}^{N}\sum_{k=1}^{L} \bar{q}_{k}\phi_{jk}^{f}\bar{q}_{j} < 1 .$$

Then all the solutions to axioms (I)-(IV) are given by

$$p_{ij} = \bar{p}_{j}\left[\frac{R_{j}^{f}R_{i}^{m}}{\sum\limits_{k=1}^{N} \bar{p}_{k}R_{k}^{f}} + \phi_{ij}^{m}\right], \qquad i = 1,\cdots,L; \quad j = 1,\cdots,N, \tag{25}$$

$$q_{ji} = \bar{q}_{i}\left[\frac{R_{i}^{m}R_{j}^{f}}{\sum\limits_{k=1}^{L} \bar{q}_{k}R_{k}^{m}} + \phi_{ji}^{f}\right], \tag{26}$$

Where the elements of the ϕ matrices have to satisfy the following set of relationships:

$$\phi_{ij}^{m} = \phi_{ji}^{f} + R_{i}^{m}R_{j}^{f}\left[\frac{\sum\limits_{k=1}^{N} \bar{p}_{k}l_{k}^{f} - \sum\limits_{k=1}^{L} \bar{q}_{k}\ell_{k}^{m}}{\left(\sum\limits_{k=1}^{L} \bar{q}_{k}R_{k}^{m}\right)\left(\sum\limits_{k=1}^{N} \bar{p}_{k}R_{k}^{f}\right)}\right]. \tag{27}$$

Special solutions, including some for male- or female-choice models can be found in Castillo-Chavez and Busenberg (1990). The following special case illustrates a model of this type that can be obtained

from 27. Let

$$\phi_{ij}^m \equiv 0, \quad \phi_{ji}^f = R_j^f \left[\frac{1}{\sum\limits_{k=1}^{N} \bar{P}_k R_k^f} - 1 \right],$$

then we obtain the following "semi-separable" mixing solution:

$$(p_{ij}, q_{ji}) = \frac{R_j^f}{\sum\limits_{k=1}^{N} \bar{P}_k R_k^f} \; (\bar{P}_j, \bar{q}_i) \; ; \tag{28}$$

where p_{ij} is independent of i, that is, males show no preference. For an alternative approach to building female (male) choice models see Section 8.

In Section 6 we will use the Ross solutions as a base model for studying the relationship between stochastic and deterministic approaches to mixing. The two-sex model described in this section is just but one way of organizing mixing sub-populations. If one wishes to model arbitrarily interacting subpopulations then one can proceed by emphasizing the interconnections between mixing subpopulations.

Demographic models that consider pairs and follow the dynamics of pairs have been studied by Kendall (1949), Fredrickson (1971), Dietz and Hadeler (1988), Dietz (1988), Hadeler (1989a,b, 1990), and Waldstätter (1989). Their approach is based on the use of a nonlinear function ψ to model the process (rate) of pair formation. This mixing/pair formation function is assumed to satisfy the Fredrickson/McFarland (1971,1972) properties:

(B1) $\psi(0,F) = \psi(M,0) = 0$

In the absence of either males or females there will be not heterosexual pair formation.

(B2) $\psi(\alpha M, \alpha F) = \alpha \psi(M,F)$ for all α, M, F $\geq$ 0.

If the sex ratio remains constant, then the increase in the rate of pair formation is assumed to be proportional to total population size.

(B3) $\psi(M + u, F + v) \geq \psi(M,F)$ for all u, v, F, M $\geq$ 0.

Increases in the number of males and/or females does not decrease the rate of pair formation.

Condition (B2) implies that all mixing functions are of the form

$$\psi(M,F) = M \, g\left(\frac{F}{M}\right) = F \, h\left(\frac{M}{F}\right),$$

where h and g are functions of one-variable.

Mixing functions satisfying the above axioms, and that have been used in demographic studies, include:

$$\psi(M,F) = k \min (M,F), \quad k \text{ is a constant}$$

$$\psi(M,F) = k \sqrt{MF} ,$$

and

$$\psi(M,F) = 2k \frac{MF}{M+F} .$$

Let σ denote the rate of pair dissolution, μ denote the natural mortality rate, Λ denote the "recruitment" rate, and W denote the number of (heterosexual) pairs. Then a simple demographic model is given by the following set of equations:

$$\frac{dM}{dt} = \Lambda - \mu M + (\sigma+\mu)W - \psi(M,F)$$

$$\frac{dF}{dt} = \Lambda - \mu F + (\sigma+\mu)W - \psi(M,F)$$

$$\frac{dW}{dt} = - (\sigma+ 2\mu)W + \psi(M,F) .$$

If Λ, μ, and σ, are constant, then there is always a globally stationary solution (M,F,W), where W is determined by the equation

$$\psi\left(\frac{\Lambda}{\mu} - W, \frac{\Lambda}{\mu} + W\right) = (\sigma+2\mu)W .$$

(for references to this and related results see Waldstätter, 1989).

Section 4. Heterogeneous sexual mixing in populations with arbitrarily connected multiple groups

A major (if largely unrecognized) drawback of the contact distribution framework is that it is very difficult to deal with incompletely connected groups (i.e., not every group mixes with every other group). If any "social" aspect to mixing is to be incorporated into contact distribution models, this deficiency must be overcome (cf Sattenspiel 1987a,b; Sattenspiel and Simon 1988; Sattenspiel and

Castillo-Chavez, 1990).

In this Section we briefly describe an extension of the heterogeneous contact distribution formalism of Busenberg and Castillo-Chavez (1989, 1990) (see also Castillo-Chavez et.al., 1990; Castillo-Chavez and Busenberg, 1990) to take account of incomplete connectance between groups and show thereby that one and two sex models, and models with multiple classes (e.g., male and female homosexuals, bisexuals and heterosexuals) can all be modeled in a unified manner, under the assumption that connectance does not change with time (see Blythe 1990 for further details and examples).

We use the same formalism of Section 1 and add the extra assumption

(vi) $0 \leq \phi_{ij} \leq U_{ij}$,   (all i, j, t)

where the $\{U_{ij}\}$ are simply the largest values of the $\{\phi_{ij}\}$ such that all the $\{R_i(t)\}$ are non-negative.

A mixing framework is said to be completely connected if every group may mix with every other group (subject to (iv)) and with itself (self-loops). With the exception of the trivial case of pure self-mixing within each group ($p_{ii}(t) = 1$, p zero elsewhere, deriving from $\phi_{ii}(t) = 1/\bar{p}_i(t)$, ϕ zero elsewhere), all solutions generated by Equation (1) are completely connected.

An incompletely connected mixing framework is one where at least one $p_{ij}(t)$ is zero for all time t, regardless of the activity levels and population sizes of the groups: people in these two groups do not mix with each other.

In principle the original formulation (Equation (1)) can handle such cases, but there are problems. Say groups l and k do not mix. Then for $p_{lk}(t) = p_{kl}(t) = 0$ for all t we require $R_l(t) = 0$ all t, and $\phi_{lk} = \phi_{kl} = 0$. Thus all the elements in the lth row and k^{th} column of p are either zeroes (as required) or else of the form $p_{ij}(t) = \bar{p}_j(t)\phi_{ij}$. As it stands, this latter form is inconvenient, as the relevant $\{\phi_{ij}\}$ need to be functions of time, such that some $R_i = 0$ while the others must satisfy $R_i \geq 0$ (essentially a linear programming problem). With just one missing connection per row, for example, all the $R_i = 0$, and we have no flexibility in the choice of the $\{\phi_{ij}\}$.

To avoid this, we re-formulate Equation (1), taking explicit account of inter-group connectedness. The re-formulated mixing framework must have the following characteristics.

First, it should permit a general description of mixing where an arbitrary number of connections are missing, and do so regardless of *which* connections these are. The second condition may be considered a corollary of the first, but is sufficiently important to be worth stating separately. If the population is partitioned into two classes, such that every group in each class mixes with every group

in the other class, but with no group in the home class, then we have a situation exactly equivalent to two-sex mixing. This case represents complete degeneracy in Equation (1). We require that our new solution be able to cope with this case, and further that it must thereby agree with the general solution to the two-sex mixing problem (for complete bipartite connectedness), recently found by Castillo-Chavez and Busenberg (1990). Of course the new solution must reduce to the Mixing formula of Busenberg and Castillo-Chavez (1989, 1990) under conditions of complete connectedness.

The third condition is at once more subtle and more fundamental. It may readily be demonstrated that not every incompletely connected mixing framework with heterogeneous $c_i(t)$ $T_i(t)$ is valid – there may not be a solution. We expect that some relatively straightforward characteristic of the new solution will reveal whether or not a solution set exists, and if so that we can converge to a member of it. Fig (4) shows a graphical representation of a simple four group population where for some $c_i(t)$ $T_i(t)$ a solution exists, but not for others.

A new solution for incompletely connected mixing groups is now presented, and its ability to meet the above conditions are evaluated in Blythe (1990). Again note that the pattern of connections between groups in assumed to hold for all time.

Borrowing from graph theory, let x be the adjacency matrix for a mixing structure, with

$$x_{ij} = \begin{cases} 1 \text{ if group i and j directly linked} \\ 0 \text{ if not} \qquad i, j = 1, 2, \cdots, N \end{cases} \tag{29}$$

Now define

$$w_{ij}(t) = \frac{x_{ij}c_j(t)\,T_j(t)}{\sum\limits_{k=1}^{N} x_{ik}c_k(t)\,T_k(t)} \tag{30}$$

i.e., $w_{ij}(t)$ is the activity of group j relative to that of all groups linked to group i, or just the *relative activity*. If groups i and j are not connected, they have relative activity of zero with respect to one another. Next, define

$$D_{ij}(t) = \frac{\sum\limits_{k=1}^{N} x_{ik}c_k(t)\,T_k(t)}{\sum\limits_{k=1}^{N} x_{jk}c_k(t)\,T_k(t)}, \qquad \text{all i and j} \tag{31}$$

as the ratio of relative activities between groups i and j. Note that for completeness we require

$$x_{ij}/x_{ji} = 0 \qquad \text{if } x_{ij} = x_{ji} = 0. \tag{32}$$

Redefining

$$R_i(t) = 1 - \sum_{k=1}^{N} w_{ik}(t)\, \phi_{ik}(t) , \qquad \text{all } i \tag{33}$$

and introducing the new quantities

$$V_i(t) = \sum_{k=1}^{N} w_{ik}(t)\, R_k(t) , \qquad \text{all } i \tag{34}$$

We may write the solution to the incompletely connected mixing problem as

$$P_{ij}(t) = w_{ij}(t)\left[\frac{R_i(t)\, R_j(t)}{V_i(t)} + \phi_{ij}(t) \right], \qquad \text{all } i \text{ and } j \tag{35}$$

which strongly resembles Equation (1). However, now the $\{\phi_{ij}(t)\}$ are no longer symmetric, as we require

$$\phi_{ji}(t) = D_{ji}(t)\phi_{ij}(t), \tag{36}$$

and $\phi_{ij}(t)$ is only defined where $x_{ij} = 1$. Note that the two-sex solutions described in the previous section are contained implicitly within Equation (35). However, the introduction of two-sex mixing functions through Axioms (I)-(IV) allows us easily to discover special solutions such as Ross solutions or the female (male) choice solutions described above. Alternatively, knowledge of the graph of the interacting subpopulations can be directly integrated into the axiomatic structure that defines the mixing matrix to generate explicit (and not obvious) special solutions.

We obtain solutions by specifying a set of ϕ_{ij} for $j \geq i$ (i.e. the upper triangular matrix), and obtain the rest ($i < j$) using constraint (iv) (note that this is arbitrary: we must specify the diagonal terms ϕ_{ii}, and then half of the remainder). Explicit conditions for the existence of solutions and several examples can be found in Blythe (1990). In the next section we present some special solutions found in the literature in the context of the framework of this section.

Section 5. Particular solutions and the general framework

Equation (1) is the general solution to the problem of describing sexual mixing among N groups within a single homosexual population, when all the groups are connected. Equation (35) permits us to relax the restriction of complete connectedness among groups. Before moving on to pair

formation/dissolution in the next section, it is worth seeing how the various published forms of mixing function fit into the general framework of (1) and (35).

(a) Random mixing. Also known as proportionate or proportional mixing, this was the original formulation of mixing by contact distribution in a heterogeneous population (*eg* Barbour 1979, Nold 1980). Proportionate mixing arises when $\phi_{ij} = K$, a constant for all i and j. From the definitions of R_i and V (Equations (3) and (4)), it is clear that this implies $p_{ij}(t) = \bar{p}_i(t)$ (Equation (2)). Regarding the elements of ϕ as measures of the inter-group preference within the population, we see that the non-uniqueness of the result (*any* K gives proportionate mixing) reflects the fact that *differences* in preference are required to give a non-random mixing pattern.

(b) Preferred mixing. This pattern of mixing is known as "assortative mating" in the population genetics literature, and takes the form

$$p_{ij}(t) = \delta_{ij}g_i + (1-g_i)\frac{(1-g_j)\bar{p}_j(t)}{\sum_{k=1}^{n}(1-g_k)\bar{p}_k(t)} , \qquad (37)$$

where the (g_i) are positive constants, and $\delta_{ij} = 1$ iff i = j, and zero elsewhere. Equation (37) has been used *e.g.* Nold (1980), Hethcote and Yorke (1984), Jacquez *et al* (1988), in the STD literature, who interpret each g_i as the fraction of partnerships of group i reserved within that group. In population genetics, g_i is interpreted as the fraction of the individulas in the group who mate exclusively among themselves. These interpretations of mixing are equivalent, but in the context of a dynamic system where disease transmission occurs, there should be different (this point is addressed by Blythe 1990). In terms of the parameters of ϕ in the general solution, we may obtain Equation (37) from

$$\phi_{ij} = \left\{ \begin{array}{ll} f_i / \bar{p}_i(t) & , \quad i = j \\ 0 & , \quad i \neq j \end{array} \right. \qquad (38)$$

(c) Like-with-like mixing. This was proposed by Blythe and Castillo-Chavez (1988), Castillo-Chavez (1989), and was generalized by Busenberg and Castillo-Chavez (1990a,b), to produce Equation (1). Here

$$\phi_{ij} = \Phi(|i-j|) \qquad (39)$$

where $\Phi(|i-j|)$ was taken to be a decreasing positive function of $|i-j|$, *ie* at a maximum for i = j, and

falling off away from the diagonal. Originally formulated for continuous variables, like-with-like mixing readily converts to the discrete-group case.

(d) "Two-phase" mixing. A number of authors have proposed particular mixing functions of the generic form (see Blythe and Castillo-Chavez 1990 for more details)

$$P_{ij}(t) = \frac{G_{ij}\bar{P}_j(t)}{\sum\limits_{k=1}^{i} G_{ik}\bar{P}_k(t)}, \tag{40}$$

The common theme is that some sexual mixing functions may be generated in two phases: first, individuals meet socially according to some underlying mixing process (in all the published cases, this has been random mixing). There is then a second phase, where one or more mutual decisions are made by individuals from groups as to the acceptability of social contacts as sexual partners,and the probability of sex actually taking place. We may relate (1) to Equation (40) by the transformation

$$G_{ij}(t) = V(t)\,\phi_{ij} + R_i(t)\,R_j(t) \tag{41}$$

so that the two-phase representation of sexual mixing may also be expressed in the general solution. Where sufficient information on social and sexual mixing and decision-making are available, this approach may be quite useful. At present, unless the restrictive assumption of *random* social mixing is made, the number of parameters required becomes much larger than ϕ, obtained by addressing the problem directly with the general solution Equation (1).

Hyman and Stanley (1988), in a continuous-activity model, refer to (the equivalent of) $G_{ij}(t)$ as "acceptance functions," symmetrical parametric functions around the $i = j$ diagonal.

Koopman et.al., (1989) consider sexual mixing, effectively within a single behavior class (see *(e)* below)according to the following schema. For a population consisting of N groups, Koopman et.al. (1989) introduce a "precursor," social, mixing process. Here all the individuals in group i have h_i social contacts per unit time, and there is some prescribed mixing pattern (Koopman et.al. (1989) use proportionate mixing). Conditional upon a social encounter between an i and a j individual, there is then a probability d_{ij} that they find each other mutually acceptable, and that they then have sex. In Koopman et.al. (1989), a component of d_{ij} is a time-dependent function which adjusts the probability of sex occurring according to the availability of prospective partners. Morris (pers. comm.) uses a simpler version of this model in her log-linear estimation scheme: all $\{h_i\}$ are assumed to be equal,

and the $\{d_{ij}\}$ are not functions.

Some of the $\{d_{ij}\}$ may have zero values for all time, if one group never accepts people from another as sexual partners, so this formulation is inherently one with incomplete connectedness. It is convenient (but does not change the model in any way) to define the usual adjacency matrix x by

$$x_{ij} = \begin{cases} 1 & \text{if } d_{ij}(t) > 0 \text{ for any } t \\ 0 & \text{if } d_{ij}(t) = 0 \text{ for all } t. \end{cases} \tag{42}$$

Then selective mixing may be written

$$P_{ij}(t) = \frac{x_{ij} \, d_{ij} \, h_j \, T_j \, (t)}{\sum\limits_{k=1}^{N} x_{ik} \, d_{ik} \, h_k \, T_k \, (t)} \tag{43}$$

Not all social encounters lead to sex, and we may calculate $\{c_i(t)\}$, the rates of acquisition of partners, per unit time, for all groups:

$$c_i(t) = \frac{h_i}{\sum\limits_{k=1}^{N} h_k T_k(t)} \sum\limits_{k=1}^{N} x_{ik} d_{ik} h_k T_k(t) \tag{44}$$

($i = 1, 2, \cdots, N$). Hence we may write

$$P_{ij}(t) = x_{ij} \, L_{ij}(t) \, c_j(t) \, T_j(t)$$

$$= w_{ij}(t) \, L_{ij}(t) \sum\limits_{k=1}^{N} x_{ik} c_k(t) T_k(t) \tag{45}$$

(all i and j), where $w_{ij}(t)$ is given by Equation (30), and

$$L_{ij}(t) = L_{ji}(t) = \frac{d_{ij} \, h_i \, h_j}{c_i(t) \, c_j(t) \sum\limits_{k=1}^{N} h_k \, T_k(t)} \tag{46}$$

Equation (43) is thus a particular case of Equation (35), where the $\{\phi_{ij}\}$ should be obtained from

$$\frac{R_i(t)R_j(t)}{V_i(t)} + \phi_{ij} = L_{ij} \sum\limits_{k=1}^{N} x_{ik} c_k T_k \tag{47}$$

In the numerical example considered by Koopman et.al. (1989), N=9, M=27, m_s=9 and r=3, so that out of a total of 36 entries in the ϕ matrix, we have at most K=30 of them available for arbitrary assignment. In fact, in this case making the ϕ some function of the L would probably be a better strategy.

It is thus clear that the general case of selective mixing falls within the circuit of Equation (35); of course, if the $\{d_{ij}\}$ are always positive, connectedness is complete and Equation (1) will suffice.

(e) "Structured mixing". The formalism of Jacquez et.al. (1989) is a complicated, parameter-rich model designed to include a variety of mixing structures for AIDS transmission modelling. We shall use slightly different notation than that of Jacquez et.al. (1989) for the sake of consistency with the rest of this paper.

The population is divided along two dimensions. The first partition is into "population sub-groups," according to type of person (e.g. drug user, male homosexual) with characteristic levels of sexual activity $\{c_j(t)\}$. The second partition is into "behavior sub-groups" (called activity sub-groups by Jacquez et.al. (1989), according to location or practices. Structured mixing may best be understood as follows.

Let $N_r(t)$ (r = 1, 2, $\cdots$, n) be the population of population sub-group r, at time t, and let $c_r(t)$ be the sexual activity (partners per unit time) of individuals in the sub-group. Now partition the members of each population sub-group according to the discrete probability density function f, such that $f_{rs}N(t)$ is the number of r-type people in behavior sub-group s (s = 1, 2, $\cdots$, m). The $\{f_{rs}\}$ are specified, and may be functions of time. Mixing is specified across R = 1, 2, $\cdots$, n within each behavior sub-group.

In order to phrase this model in the language of this paper, we must "unpack" the compact notation of Jacquez et.al. (1989). We regard the behavior groups as labelled blocks of an (n $\times$ m) $\times$ (n $\times$ m) mixing matrix. Each such block is comprised of n rows (one for each level of activity) and n columns (for those portions of mixing occurring within the behavior group). The $\{c_j(t)\}$ in the mixing matrix are reflected by

$$c_{\ell+km} = c_\ell \quad \ell = 1, 2, \cdots, m \quad k = 1, 2, \cdots, n \tag{48}$$

reflecting the fact that the same population sub-groups are represented in every behavior sub-group. The population associated with each class (i.e. each row of our block-composed mixing matrix) is given by

$$T_{r+(s-1)m} = f_{rs}N_r(t) \quad r = 1, 2, \cdots, n \quad s = 1, 2, \cdots, m \quad (49)$$

The adjacency matrix x contains no zeroes: absence of contact between classes i and j (say) occurs because one of these classes is empty (some $f_{rs} = 0$). This is covered by constraint (iv), so that structured mixing is $\underline{not}$ an example of incomplete connection, and is covered by Equation (1) (with one provision discussed in the next example).

It should be noted that Jacquez et.al. (1989) do not introduce the matrix $\underline{f}$ with the interpretation (partition of r-people across the s-groups) used here. They interpret f_{rs} as the fraction of the partners of a r-person who come from behavior group s. The two interpretations are equivalent, as f_{rs} acts as a scale factor for N_r in behavior groups (c.f. Jacquez et.al., 1989, p. 310); this may clearly be seen by writing the balance constraint (iii) in the manner of Jacquez et.al. (1989) for sexual contacts between group r and group r, individuals in behavior group s:

$$c_r N_r f_{rs}\rho(s)_{rr\prime} = c_{r\prime}N_{r\prime}f_{r\prime s}\rho(s)_{r\prime r} \quad (50)$$

Note that if an incompletely connected mixing model is used within one or more of the behavior groups of Jacquez et.al. (1989), then of course the structured mixing model is incompletely connected, and Equation (35) rather than Equation (1) must be used. The difference between zeroes in $\{d_{ij}\}$ and zeroes in $\{f_{rs}\}$ should be appreciated.

These and other examples are considered in Blythe and Castillo-Chavez (1990) and Blythe (1990).

Section 6. Stochastic Simulation Mixing Models

We have addressed the issue of stochastic effects in HIV transmission dynamics using three separate approaches: stochastic simulations; probability evolution models; and stochastic differential equations. So far we have used stochastic simulations to study the relationships between stochastic pair formation models, deterministic pair formation models, and deterministic contact distribution models. The probability evolution models and stochastic differential equations have been applied in an attempt to see how well (or badly) deterministic models and *approximate* stochastic models do when group sizes are small. We provide a brief overview of the stochastic simulation work below.

We have developed a simulation program, running under FORTRAN or PASCAL on either SPARC-1 workstations or the Cornell IBM supercomputer, depending on problem size. A large population of individuals (classified in any specified manner, eg. by age, sex, sexual orientation or disease status) can be followed through the probabilistic sequence of pair formations and dissolutions,

with preferences (in the form of enhanced formation probabilities, for example) an integral feature. As yet we have performed no systematic epidemiological investigations using this technology, but instead have concentrated on understanding the relationships between stochastic pair formation and deterministic models.

For example, say we follow a group of $N=10^3$ "homosexuals" for sufficient time to achieve stationary behavior, over numerous replicates (eg. 10^2). If we have introduced no preference structure, then the results of the random mixing should be comparable with the proportionate mixing assumptions (see Equations 1 and 2) underlying much deterministic contact distribution modelling. Figure 5 illustrates that this is indeed the case.

Likewise, we can look at non-preferential pair formation/dissolution simulations with two-sexes, and compare the results with the expected Ross's solutions. Figure 6 again illustrates that the deterministic version of heterosexual "random mixing" is in agreement with stochastic descriptions.

We may extend this to more interesting scenarios, where we posit preference rules among types of individuals in a two-sex model, and look at the p_{ij} and q_{ij} surfaces that arise from the simulation. Figure 7 shows two such surfaces, ranging from Ross's (random) solutions to a case where individuals strongly preferred partners (of the other sex) who were as sexually experienced as they were. Characterizing the strength of "like-with-like" association by the correlation coefficient for the data in such figures, we find a systematic "linear" relationship between correlation and Q (an inverse measure of preference) over much of the range (Fig 8). We are presently developing approaches for comparing group preferences (deterministic models) versus individual preferences (stochastic models) so as to obtain similar "limiting" results as those corresponding to Ross solutions (random heterosexual mating). In the next section we discuss the potential use of these mixing functions in the modeling of genetic mating systems.

Section 7. Genetics models with preferential mating

It is clearly of interest to see whether these new descriptions of mixing have any relevance in population genetics where mating systems play a fundamental role. Here we discuss our preliminary results applying the general N-group mixing function of Busenberg and Castillo-Chavez (1989, 1990) to an old and very simple problem; namely, the effects of preferential mating on genotype frequency for a single locus/recessive allele system.

A description has already been provided in Section 1 through Axioms (i)-(iv) with Equations (1)-

(5) providing a useful characterization of all p matrices. Equation (1) was formulated for homosexual (i.e. one-sex) mixing, so its applicability to sexual reproduction is of course limited. Here we begin by using the simpler formalism, summarized in (1), as an approximation ("pseudo-sexual mating" for monoecius populations) to the equal sex-ratio large population case. We are aware of the pitfalls for the unwary of neglecting explicit sexual reproduction.

As a first attempt at applying the generalized mixing framework given by Equation (1) to population genetics, let us consider an elementary text-book example, specifically that of a recessive gene at a single locus, with no population regulation or frequency dependence. A good example (Crow 1986, pp. 50-53) is red-headedness in a human population. We divide the population into three groups: [1] homozygous (AA) individuals who do not carry the "red" allele; [2] heterozygotes (Aa) who carry but do not express the allele; and [3] homozygotes (aa) who express red hair. We may assume that the fractions of offspring born from the six possible crosses follows the elementary theory (see Crow, 1986).

Clearly no one can distinguish AA from Aa phenotypes, so the only reasonable form for the preference matrix is

$$\phi = \begin{bmatrix} \alpha & \alpha & \beta \\ \alpha & \alpha & \beta \\ \beta & \beta & \gamma \end{bmatrix} \tag{51}$$

where $0 \leq \alpha, \beta, \gamma \leq 1$ are constants. Individuals in group 1 and 2 have the same preferences for aa versus non-aa (β versus α), and group 3 individuals have preferencies β and γ for non-aa and aa individuals, respectively.

We should note at this point that assortative mating is represented by $\beta = \alpha = 0$, and $\gamma = r/\bar{p}_3(t)$ where $r < 1$. The result that a fixed fraction of the population of aa individuals mate among themselves, regardless of group population sizes, has attracted considerable criticism in the field of sexually-transmitted disease epidemiology, and does not seem to be biologically reasonable (see Castillo-Chavez 1989, Castillo-Chavez et al. 1989a, b, Huang et al. 1990).

With $\alpha = \beta = \gamma$ taking any value in [0,1], the mixing structure reduces to proportionate or random mixing. For the pseudo-sexual model presented here, we make the following standard simplifying assumptions:

(a)   Every individual in generation n has just one partner (mate).

(b)   The unit of time is the generation.

(c)   Individuals from generation n are not counted in generation n+1.

(d)   All matings produce 2b offspring (b>0).

(e)   The a allele frequency q remains constant.

Then, $c_i(t) = 1$, all i and t, and

$$\bar{P}_i(t) = \frac{T_i(t)}{T(t)} \equiv x_i(t) \text{ for all } i, \ T(t) = \sum_{k=1}^{N} T_i(t), \tag{52}$$

where $T(t)$ is the total population, and $x_i(t)$ is the proportion of group i in the population, in generation t. We have

$$R = \begin{bmatrix} 1 - \alpha(x_1+x_2) - \beta x_3 \\ 1 - \alpha(x_1+x_2) - \beta x_3 \\ 1 - \beta(x_1+x_2) - \gamma x_3 \end{bmatrix}$$

$$= \begin{bmatrix} (1-\alpha) + (\alpha-\beta) x_3 \\ (1-\alpha) + (\alpha-\beta) x_3 \\ (1-\beta) + (\beta-\gamma) x_3 \end{bmatrix} \tag{53}$$

and

$$V = (1-\alpha) + 2(\alpha-\beta)x_3 - (\alpha-2\beta+\gamma)x_3^2. \tag{54}$$

Note that $x_1 + x_2 + x_3 = 1$ and $\frac{1}{2} x_2 + x_3 = q$.

We now construct the encounter matrix **M**, where $M_{ij} = T_i p_{ij}$ represents the total number of pairs formed by group i individuals with group j individuals in a given generation. Each pair produces 2b offspring, so there are bT(t) offspring in generation t+1. The fraction of offspring who are aa is seen to be

$$x_3(t+1) = \frac{1}{T(t)}\left[\frac{1}{4} M_{2,2}(t) + \frac{1}{2} M_{2,3}(t) + \frac{1}{2} M_{3,2}(t) + M_{3,3}(t)\right] \tag{55}$$

(half the "cross" offspring are counted as 3 × 2 , and half as 2 × 3).

Now for convenience write $Z_t = x_3(t)$ as the proportion of aa individuals in the population in

generation t. Then we find the recurrence relation

$$Z_{t+1} = f(Z_t) + g(Z_t)^2 / h(Z_t), \qquad Z_0 < q \tag{56}$$

where

$$f(Z) = \alpha q^2 - 2(\alpha-\beta)qZ + (\alpha-2\beta+\gamma)Z^2$$
$$g(Z) = (1-\alpha)q + (\alpha-\beta)(1+q)Z - (\alpha-2\beta+\gamma)Z^2 \tag{57}$$
$$h(Z) = (1-\alpha) + (\alpha-\beta)Z - (\alpha-2\beta+\gamma)Z^2.$$

Note that we may recover the familiar assortative mating result with $\alpha = \beta = 0$, $\gamma = r/Z$, i.e.

$$Z_{t+1} = rZ_t + (q-rZ_t)^2 / (1-rZ_t). \tag{58}$$

Fig (9) illustrates two solutions to (56) for some arbitrary parameter values. A result appreciably different from that produced by random mixing (*i.e.* $\alpha = \beta = \gamma$) requires a fairly strong tendency to avoid mixing outside the home group; the behavior of the largest group ($AA + Aa$) dominates.

Section 8. Modeling Female Choice

Any modeling exercise involves compromises between detail and tractability, and requires that, somewhere along the line, particular functional forms approximating our knowledge of particular processes be chosen. In addition, of course, values (or ranges of values) of the model parameters must be arrived at, by a process ranging from sheer guesswork to rigorous statistical estimation. Modeling pair formation/dissolution in human populations is no exception to the above, and in many ways, due to the extreme paucity of our knowledge about human social and sexual decision-making in general, we find ourselves in the worst of all possible worlds for a modeler. Any remotely realistic model tends to be highly detailed, with an embarrassing proliferation of parameters and functions about which we know next to nothing.

Busenberg and Castillo-Chavez (1989, 1990) have provided us with a general framework for pair formation/dissolution in a heterogeneous age-structured heterosexual population, and indicated how this may be used in studying sexually transmitted disease dynamics. There remains the task, however, of converting their general framework into practical schemes for modeling the pair formation/dissolution processes in a given modeling context. The following is a strategy for doing so.

1. We must choose the number of groups comprising the female (L groups) and the male (N

groups) sub-populations. Heterogeneity requires at least two of each. If N and L are large, the group populations will be in most cases small, and the rates of pair formation between any two smaller still. This implies that it will be difficult perhaps even to detect some groups in the population, using standard survey methods, and that estimation of inter-group pairing may be all but impossible for such poorly represented groups. Clearly we must compromise; from direct examination of surveys of sexual behavior, and from experimentation with models, we suspect that 4 or 5 represents an upper feasible limit on N and on L. What the groups represent must of course depend on the particular system being modeled.

2. The choice of male-female affinities or preferences for the groups (the ϕ matrices) represents what is probably the greatest unknown in the modeling process. Estimation schemes for ϕ are being developed (Blythe *et al* 1990), but will always be severely impeded by the difficulty of obtaining the reliable serial data on mixing that such schemes seem to require. Models may be used to investigate the impact of different ϕ-structures on the pair formation/dissolution process, with the hope of gaining insight into the dynamics of these processes in real populations, and perhaps as a consequence restricting the ϕ-structure to a particular structure or structures for human populations. The actual choices of ϕ must be dependent upon the particular population (or type of population) in which we are interested; for example, if we include female prostitutes as a group, we would expect a radically different structure than if only non-professionals are present. Parametric ϕ-generating functions may be a good place to start in the investigation of affinity structures. Although it is possible to construct schemes where the ϕ for males and the ϕ for females are essentially uncorrelated (*e.g.* the "semi-separate" solutions Eq (29)), a simple but effective mixing function may be geneerated by making the plausible assumption

$$\phi_{ij}{}^m = \phi_{ji}{}^f \quad , \qquad i = 1, 2, \cdots, N; \; j = 1, 2, \cdots, L$$

3. Conservation of pair formation rates. The Busenberg and Castillo-Chavez (1989, 1990) framework incorporates rates of partner acquisition (per unit of time) for each male and each female group in the population. However, it is clear that not all of these can be constants, as the total rate of pair formation in the population must be conserved at all times. We may express this constraint thus:

$$\sum_{k=1}^{N} b_k T_k^m(t) = \sum_{\ell=1}^{L} c_\ell T_\ell^f(t) \qquad , \text{all } t \geq 0 \tag{59}$$

(c_i and $T_i{}^m$ are respectively the per capita partner acquisition rate and group population size for male group i; b_j and $T_j{}^f$ are the equivalent quantities for female group j). Clearly, a part of the process of

model formulation must include an assumption about which acquisition rates remain constant, and which must adjust to satisfy the constraint. Of course, there are an infinite number of ways to do this, but that's not the point. It should be realized that we have to choose which groups are dominant (in the sense that they achieve their target partner acquisition rates), which are sub-dominant (they must adjust their rates to a level set by the constraint), and how the adjustment of sub-dominant acquisition rates may be described. This can only be done for a given system, or type of system, where we are armed with information or assumptions concerning the behavior and dominance of groups.

A good place to start would be with one sex having a consistent degree of "formation-dominance" over the other. Then if we define β_i^m (β_j^f) as the value of partner acquisition rate c_i (b_j) under conditions where the female (male) groups are completely dominant -- i.e. b_i (c_j) is maintained -- we may approximate the group acquisition rates by

$$\dot{c}_i = (1-\delta)c_i + \delta\beta_i^m \quad , \qquad i = 1, 2, \cdots, N \tag{60}$$

$$\dot{b}_j = \delta b_j + (1-\delta)\beta_j^f \quad , \qquad j = 1, 2, \cdots, L \tag{61}$$

where $0 < \delta < 1$ is the *degree of relative female dominance in the pair formation process* i.e., $\delta = 1$ implies females are completely dominant, and males adjust, while $\delta = 0$ implies the converse.

We have still the modeling task of choosing how to calculate the β_i^m and β_j^f. Again this should really be done in the light of behavioral information. A plausible first approximation at a general scheme, however, might be to have the β's proportional to the level of affinity or preference for the group in the (other sex) population, and inversely proportional to the representation of its group in the (home sex) population. The implication is that "popular" groups tend to have higher acquisition rates, and that being scarce increases the rate for *individuals* in any group. If we define

$$n_i^m = \frac{T_j^m}{\sum\limits_{n=1}^{N} T_k^m} \, , \qquad n_j^f = \frac{T_j^f}{\sum\limits_{l=1}^{L} T_k^f} \tag{62}$$

as the fractional contributions of particular groups to the population of the same sex, and

$$\overline{\phi}_i^m = \sum\limits_{l=1}^{L} n_l^f \phi_{l_i}^f \, , \qquad \overline{\phi}_j^f = \sum\limits_{k=1}^{N} n_k^m \phi_{k_j}^m \tag{63}$$

as the (weighted) average affinity or preference levels for each group, then we could define the "totally sub-dominant" acquisition rates according to

$$\beta_i^m = \frac{1}{T_i^m} \sum_{l=1}^{L} c_l T_l^f \frac{\overline{\phi}_i^m}{\sum\limits_{k=1}^{N} \overline{\phi}_k^m}, \tag{64}$$

$$\beta_j^f = \frac{1}{T_j^f} \sum_{k=1}^{N} b_k T_k^m \frac{\overline{\phi}_j^f}{\sum\limits_{l=1}^{L} \overline{\phi}_l^f} . \tag{66}$$

4. This gives us a complete model specification for the pair formation/dissolution process: the general framework, plus specific affinity, dominance, and adjustment features. We may now, for example, study the effect of relative dominance on pair formation/dissolution by varying δ between 0 and 1. Of course, we must have specified all the parameters for affinity, recruitment, loss rates, pair dissolution rates, "target" acquisition rates, and initial rates for T_i^m and T_j^f — as many as $4 \times N \times L + 5 \times (N+L)$ parameters (i.e., for $N = L = 4$, we have up to 104 parameters, if individuals between groups are different and behave differently in all the possible pairwise combinations). Clearly, we will want to simplify as much as possible, but only appropriate data can give us safe justification for doing so!

Section 9. Discussion

In this paper we have illustrated a very general approach to the modeling of the processes of pair-formation and dissolution. The flexibility of Equation (1) has been demonstrated in a variety of biological areas including demography, epidemiology, and genetics, but the work has just begun. What are the implications of heterogenously mixing populations in the dynamics of sexually-transmitted diseases, is still an open question. However, Equation (1) provides us with a way of approaching this problem in a systematic manner (see Blythe and Castillo-Chavez 1990b, Castillo-Chavez and Blythe 1990). We finally note that the above process can be used in the context of age-structured populations. We have limited ourselves to the non-age structured versions to make our formulae simpler making it more accesible to a larger group of biologists or epidemiologists.

REFERENCES(needs to be updated)

Anderson, R. M. and May, R. M. (1984). Spatial, temporal and genetic heterogeneity in host populations and the design of immunization programmes. *IMA J. of Math. Applied in Med. and Biol.* 1: 233-266.

Anderson, R. M., S. P. Blythe, S. Gupta, and E. Konings. (1989). The transmission dynamics of the Human Immunodeficiency Virus Type 1 in the male homosexual community in the United Kindom: the influence of changes in sexual behavior. (Manuscript.)

Barbour, A. D. (1978). MacDonald's model and the transmission of bilharzia. *Trans. Royal Soc. Trop. Med. Hyg.* 72: 6-15.

Blythe, S. P. (1990). Heterogeneous sexual mixing in populations with arbitrarily connected multiple groups (Manuscript.).

Blythe, S.P. and R.M. Anderson. (1988b). Variable infectiousness in HIV transmission models. *IMA J. of Mathematics Applied in Med. and Biol.* 5, 181-200.

Blythe, S.P. and C. Castillo-Chavez. (1989). Like-with-like preference and sexual mixing models. *Math. Biosci.* 96, 221-238.

Blythe, S.P. and C. Castillo-Chavez. (1990a). "The one-sex mixing problem: a choice of solutions? (Submitted *J. of Math. Biol.*)

Blythe, S.P. and C. Castillo-Chavez. (1990b). Like-with-like mixing and sexually transmitted disease epidemics in one-sex populations (Manuscript.)

Blythe, S.P., C. Castillo-Chavez, and G. Casella. (1990). Empirical methods for the estimation of the mixing probabilities for socially structured populations from a single survey sample. (Submitted *Math. Pop. Studies.*)

Busenberg, S. and C. Castillo-Chavez. (1989). Interaction, pair formation and force of infection terms in sexually transmitted diseases. In (C. Castillo-Chavez, ed.) *Mathematical and Statistical*

Approaches to AIDS Epidemiology. Lecture Notes in Biomathematics 83, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong: 289-300

Busenberg, S. and C. Castillo-Chavez. (1990). A general solution of the problem of mixing of subpopulations, and, its application to risk- and age- structured epidemic models. (To appear in *IMA J. of Math. Applic. to Med. and Biol.*)

Castillo-Chavez, C. (1989). Review of recent models of HIV/AIDS transmission. In (S. A. Levin, T. G. Hallam, and L. J. Gross, eds.) *Applied Mathematical Ecology*, Biomathematics 18, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong: 253-262

Castillo-Chavez, C. and S. P. Blythe. (1989). Mixing framework for social/sexual behavior. In (Castillo-Chavez, ed.) *Mathematical and Statistical Approaches to AIDS Epidemiology.* Lecture Notes in Biomathematics 83, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong: 275-288

Castillo-Chavez, C. and S.P. Blythe. (1990). A "test-bed" procedure for evaluating one-sex mixing frameworks (Manuscript.)

Castillo-Chavez C. and Busenberg S. (1990). On the solution of the two-sex problem (submitted.)

Castillo-Chavez, C., Hethcote, H., Andreasen, V., Levin, S. A., and Liu, W-m. (1988.) Cross-immunity in the dynamics of homogeneous and heterogeneous populations. In: *Mathematical Ecology* (T. G. Hallam, L. G. Gross, and S. A. Levin, eds.). pp. 303-316. World Scientific Publishing Co., Singapore.

Castillo-Chavez, C., Hethcote, H., Andreasen, V., Levin, S. A., and Liu, W-m. (1989). Epidemiological models with age structure, proportionate mixing, and cross-immunity. *J. Math. Biol.*, 27(3): 233-258.

Castillo-Chavez, C., Cooke, K., Huang, W. and Levin, S. A. (1989a). On the role of long incubation periods in the dynamics of HIV/AIDS. Part 2: Multiple group models. In: *Mathematical and statistical approaches to AIDS epidemiology* (C. Castillo-Chavez, ed.), pp. 200-217. Lecture Notes in Biomathematics 83, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong.

Castillo-Chavez, C., Cooke, K., Huang, W. and Levin, S. A. (1989b). Results on the dynamics for models for the sexual transmission of the human immunodeficiency virus. *Applied Math. Letters*, 2(4): 327-33L

Castillo-Chavez C, Busenberg S, Gerow K. (1990a). Pair formation in structured populations. *Proceedings of International Conference on Differential Equations and Applications* (W. Schappacher, ed.), Retzhof, Austria 1989. (to appear.)

Castillo-Chavez, C., Palmer, J. S., and S. P. Blythe (1990b). Assortative versus Non-Assortative Pair-formation: the case of two interacting homosexually active populations (manuscript).

Cooke, K. L. and J. A. Yorke. (1973). Some equations modelling growth processes and gonorrhea epidemics. *Math. Biosci.*,58, 93-109

Dietz, K. (1988). On the transmission dynamics of HIV. *Math. Biosci.* 90, 397-414.

Dietz, K. and K.P. Hadeler. (1988). Epidemiological models for sexually transmitted diseases. *J. Math. Biol.* 26, 1-25.

Dietz, K. and Schenzle, D. (1985). Proportionate mixing models for age-dependent infection transmition. *J. Math. Biol.* 22, 117-120.

Fredrickson, A.G. (1971). A mathematical theory of age structure in sexual populations: Random mating and monogamous marriage models. *Math. Biosci.* 20, 117-143.

Gimelfarb A. (1988a). Processes of pair formation leading to assortative mating in biological populations: encounter mating model. *The American Naturalist.* 131 (6): 865-884.

Gimelfarb A. (1988b). Processes of pair formation leading to assortative mating in biological populations: dynamic interaction model. *Theor. Pop. Biol.* 34: 1-23.

Hadeler, K.P. (1989a). Pair formation in age-structured populations. *Acta Applicandae Mathematicae* 14, 91-102.

Hadeler, K.P. (1989b). Modeling AIDS in structured populations. (Manuscript.)

Hadeler, K.P. (1990). Homogeneous delay equations and models for pair formation. (Manuscript.)

Hethcote, H.W. and J.A. Yorke. (1984). *Gonorrhea transmission dynamics and control.* Lecture Notes in Biomathematics 56, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

Huang, W. (1989). Studies in differential equations and applications. Ph. D. Thessis, The Claremont Graduate School (December 1989), Claremont CA.

Huang, W., K. Cooke and C. Castillo-Chavez. (1989). Stability and bifurcation for a multiple group model for the dynamics of HIV/AIDS. (To appear in *SIAM J. of Applied Mathematics*).

Hyman, J.M. and E.A. Stanley. (1988). Using mathematical models to understand the AIDS epidemic. *Math. Biosci.* 90, 415-473.

Hyman, J.M. and E.A. Stanley. (1989). The effect of social mixing patterns on the spread of AIDS. In *Mathematical approaches to problems in resource management and epidemiology.* C. Castillo-Chavez, S.A. Levin, and C. Shoemaker (eds.). Lecture Notes in Biomathematics 81, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo: 190-219

Jacquez, J.A., C.P. Simon, J. Koopman, L. Sattenspiel, and T. Perry. (1988). Modelling and analyzing HIV transmission: the effect of contact patterns. *Math. Biosci.* 92, 119-199.

Jacquez, J.A., Simon, C.P. and Koopman, J. Structured mixing: heterogeneous mixing by the definition of mixing groups. *Mathematical and Statistical Approaches to AIDS Epidemiology* (C. Castillo-Chavez, ed.) Lecture Notes in Biomathematics 83, 301-315. Springer-Verlag (1989).

Kendall, D.G. (1949). Stochastic processes and population growth. *Roy. Statist. Soc., Ser. B* 2, 230-264.

Keyfitz, N. (1949). The mathematics of sex and marriage. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability.* Vol. IV: Biology and Health, 89-108.

Koopman, J.S., Simon, C.P., Jacquez, J.A. and Tae Sung Park (1989). Selective contact within structured mixing with an application to HIV transmission risk from oral and anal sex. In: *Mathematical and statistical approaches to AIDS epidemiology* (C. Castillo-Chavez, ed.), pp. 316-348. Lecture Notes in Biomathematics 83, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong.

McFarland, D. D. (1972). Comparison of alternative marriage models. In (Greville, T. N. E., ed.), *Population Dynamics.* Academic Press, New York London: 89-106.

Nold, A. (1980). Heterogeneity in disease-transmission modeling. *Math. Biosci.* 52, 227-240.

Parlett, B. (1972). Can there be a marriage function? In (Greville, T. N. E., ed.), *Population Dynamics.* Academic Press, New York London: 107-135.

Pollard, J. H. (1973). *Mathematical models for the growth of human populations,* Chapter 7: The two sex problem. Cambridge University Press.

Thieme, H. R. and Castillo-Chavez, C. (1989). On the role of variable infectivity in the dynamics of the human immunodeficiency virus epidemic. In: *Mathematical and statistical approaches to AIDS epidemiology* (C. Castillo-Chavez, ed.), pp. 157-176. Lecture Notes in Biomathematics 83, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong.

Thieme, H. R. and Castillo-Chavez, C. (1990). On the possible effects of infection-age-dependent infectivity in the dynamics of HIV/AIDS. Biometrics Unit Technical Report, BU-1102-M.

Waldstätter, R. (1989). Pair formation in sexually transmitted diseases. In (C. Castillo-Chavez, ed.) *Mathematical and Statistical Approaches to AIDS Epidemiology.* Lecture Notes in Biomathematics 83, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong: 260-274

Sattenspiel, L. (1987a). Population structure and the spread of disease. *Human Biol.* 59, 411-438.

Sattenspiel, L. (1987b). Epidemics in non-randomly mixing: a simulation. *Am. J. Physics, Anthropol.* 73, 251-265

Sattenspiel, L. and C. Castillo-Chavez. (1990). Environmental context, social interactions and the spread of HIV. *Amer. J. Human Biology* 2 (4). (In press.)

Sattenspiel, L. and C.P. Simon (1988). The spread and persistence of infectious disease in structured populations. *Math. Biosci.* 90, 341-366

Figure Captions:

Fig 1. Surfaces of $p_{ij}(t)$ from the SIS model (6)-(8) with parameters: $A = (5,4,3,2,1)$ individuals per year; $c = (1,2,3,8,20)$ partners per year; $T(0) = (1000,500,250,125,62)$; $\mu = 0.01$ per year; $a = 0.8$ and $b = 0.2$ (like-with-like mixing). *(a)* $p_{ij}(25)$, early in the transient behavior of the system, and *(b)* $p_{ij}(500)$, when a steady state has been very closely approached.

Fig 2. Illustration of how the general mixing result Eq (1) may be represented as the result of two processes: first the preferred reservation of partners, and then a random apportionment of the remainder. Although real mixing processes probably never operate this way, we can always describe them as if they had.

Fig 3. Shematic outline of the pair formation/dissolution SIS model of Section 3. Key: f = single uninfected females, F = single infected females; m = single uninfected males; M = single infected males; fm = uninfected pair; fM = uninfected female and infected male in pair; Fm = infected female and uninfected male in pair; FM = both partners infected in pair.

Fig 4.

1. Not all mixing problems have a solution. Shown are two graphs representing N=4 mixing cases and their adjacency matrices. Each circle represents a group (number in parenthesis is group number) with a given population and characteristic activity level (lower number in each circle, c_i T_i). Lines between groups indicate that two groups mix, and self-loops join a group to itself. (a) No solution possible (b) An infinite number of solutions possible.

Fig 5. Comparison between stochastic and determinisitic results for a homosexual partnership process. The stochastic simulation model considers pair formation/dissolution according to specified distributions of individual probabilities per time-step of joining with a *randomly* encountered prospective partner, and of initiating pair dissolution if paired. The determinsitic analogue is the random mixing solution $p_{ij}(t) = \bar{p}_j(t)$ from Eq (2). The groups are calculated for large t by the number of partners individuals have had on average per time-step (*i.e.* the average c_i). If the two curves coincide exactly, then the determinisitic contact-distribution model captures the behavior of the stochastic pair-formation model. For this random mixing case, agreement is seen to be very good.

Fig 6. Comparison between stochastic and deterministic results for a two-sex partnership process. The determinisitic analogue is here composed of the Ross solutions (Eqns (25) and (26), and Theorem 1 of Section 3), using the c_i and T_i from the stochastic model at large t when the process approximates stationarity. The stochastic simulation model is the two-sex equivalent of that for the homosexual process, with different distributions of individual pair formation and dissolution-initiation probabilities per unit time for each sex. Displayed is the comparision of the nominal females; exact coincidence of curves would imply that the deterministic description captures the behavior of the simulation. For this random mixing example, agreement is seen to be very good.

Fig 7. Behavior of the two-sex stochastic simulation model when there is non-zero preference betweem groups. Shown is the joint distribution of sexual partnerships across all male and female groups. Preference in the simulation was characterised by a parameter Q which influenced the probability of two individuals forming a pair based on the difference between their sexual activiies (numbers of partners per unit time); this is clearly a stochastic pair analogue of the "like-with-like" determinisitic mixing scenario. As $Q \rightarrow \infty$ the pure random mixing (Ross solution) is rapidly approached, while for $Q \rightarrow 0$ individuals acceept only partners of the same activity level. (a) $Q = 1.0$; already the shape of the distribution suggets almost random mixing. (b) $Q = 1.0 \times 10^{-7}$; an extreme preference case, where the diagonal is the dominant feature.

Fig 8. For a series of simulation runs such as those illustrated in Fig 6, we can calculate the correlation coefficient (r) associated with a joint distribution of female/male partnerships. Plotting r against $\log_{10} Q$, we find that that there is a "scaling region" of preferences Q where $r \propto -\log_{10} Q$.

Fig 9. Examples of the "red-headed" monoecious model Eq (56), with $q = 0.1$, $Z_0 = 0.005$. *In-mixing:* here $\alpha = 1.0$, $\beta = 0.0$, $\gamma = 1.0$. *Out-mixing:* here $\alpha = 0.0$, $\beta = 1.0$, $\gamma = 0.0$. The out-mixing case does not differ appreciably from the "random" case $\alpha = \beta = \gamma = $ constant, where Hardy-Weinberg values are reached in one generation. Strong in-mixing, in particular among the largest group (the *AA* and *Aa people*), is required for the slower convergence shown.
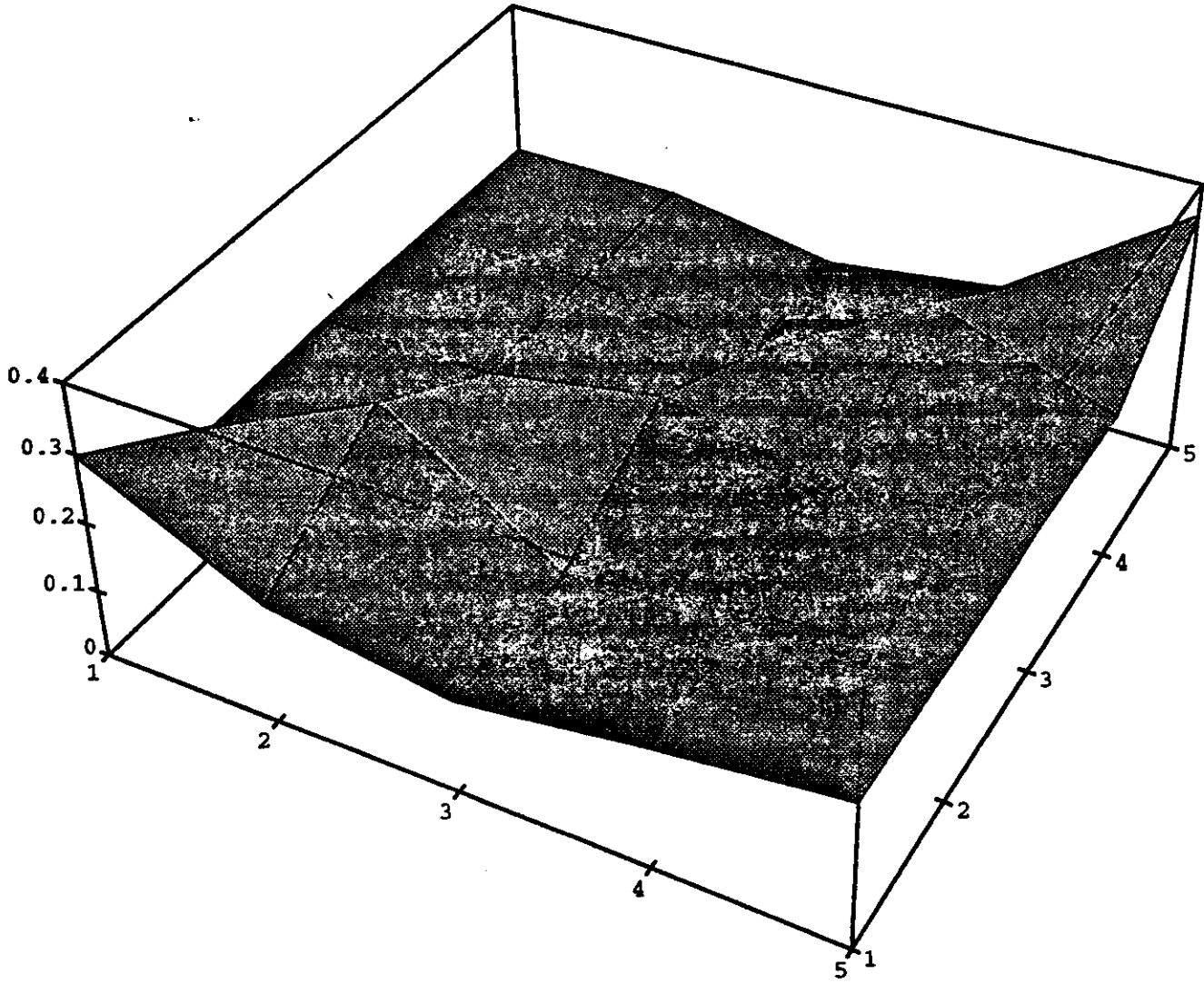
Fig 1(a) Liho thir
$t=0$
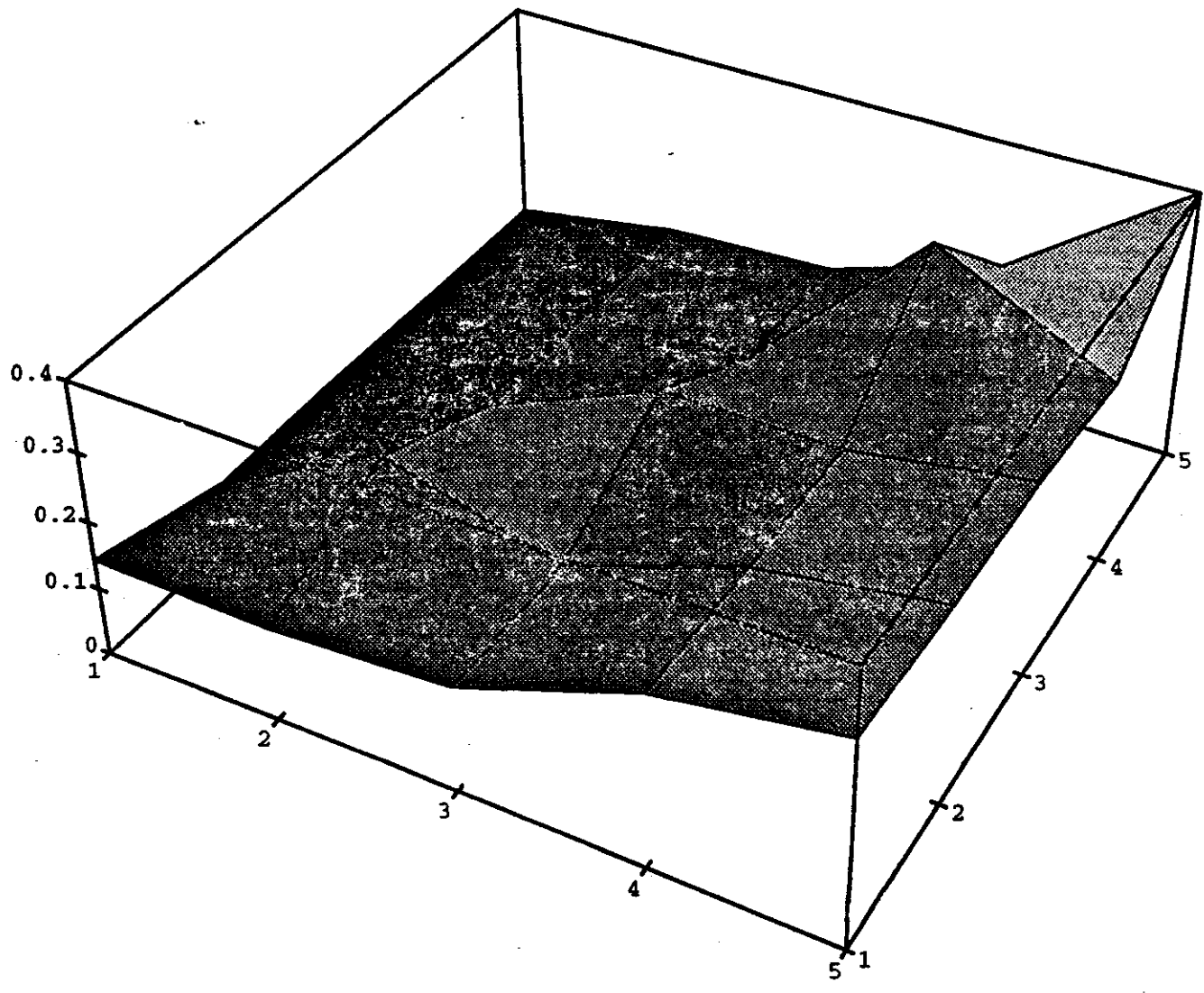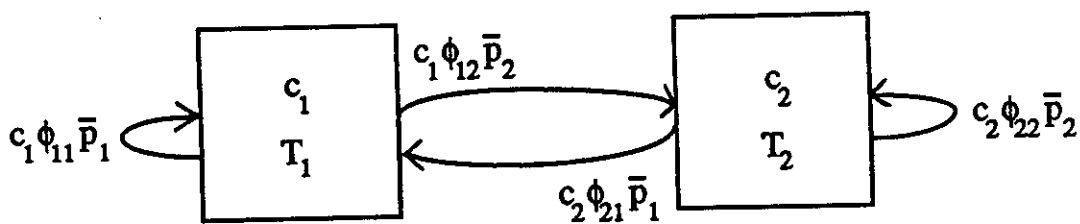
$(ob1 = (.8, .2)$

Fig 1(b)   t = 500

as = (0.8, 0.2)

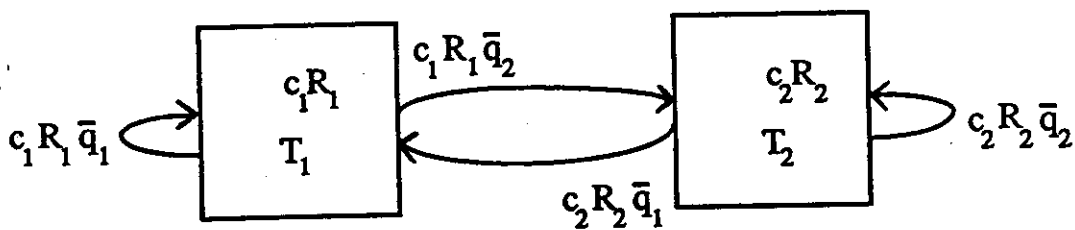$$\text{Fig 2}$$

PREFERENCE

$$\boxed{\begin{array}{c} c_1 \\ T_1 \end{array}} \qquad \boxed{\begin{array}{c} c_2 \\ T_2 \end{array}}$$

$c_1 \phi_{11} \bar{P}_1$

$c_1 \phi_{12} \bar{P}_2$

$c_2 \phi_{21} \bar{P}_1$

$c_2 \phi_{22} \bar{P}_2$

$\oplus$

PROPORTIONATE MIXING

$$\boxed{\begin{array}{c} c_1 R_1 \\ T_1 \end{array}} \qquad \boxed{\begin{array}{c} c_2 R_2 \\ T_2 \end{array}}$$

$c_1 R_1 \bar{q}_1$

$c_1 R_1 \bar{q}_2$

$c_2 R_2 \bar{q}_1$

$c_2 R_2 \bar{q}_2$

$$\bar{q}_1 = \frac{c_1 R_1 T_1}{c_1 R_1 T_1 + c_2 R_2 T_2} \qquad \bar{q}_2 = \frac{c_2 R_2 T_2}{c_1 R_1 T_1 + c_2 R_2 T_2}$$

$\Downarrow$

GENERAL MIXING

37

Fig 3

(a)
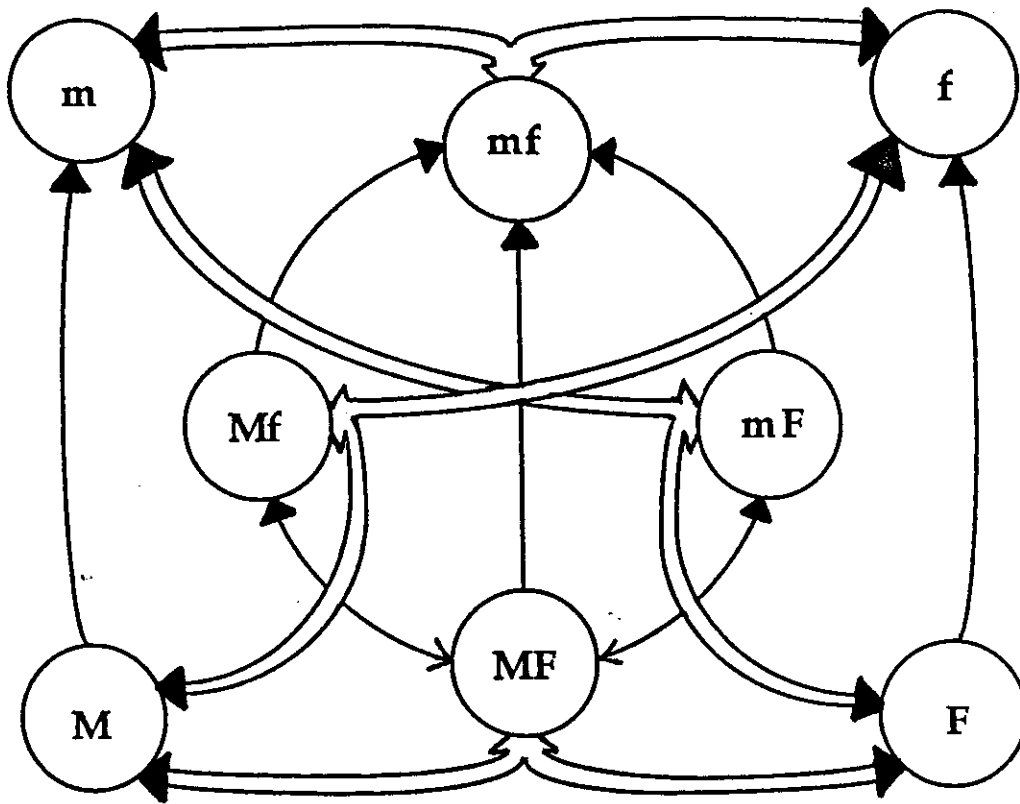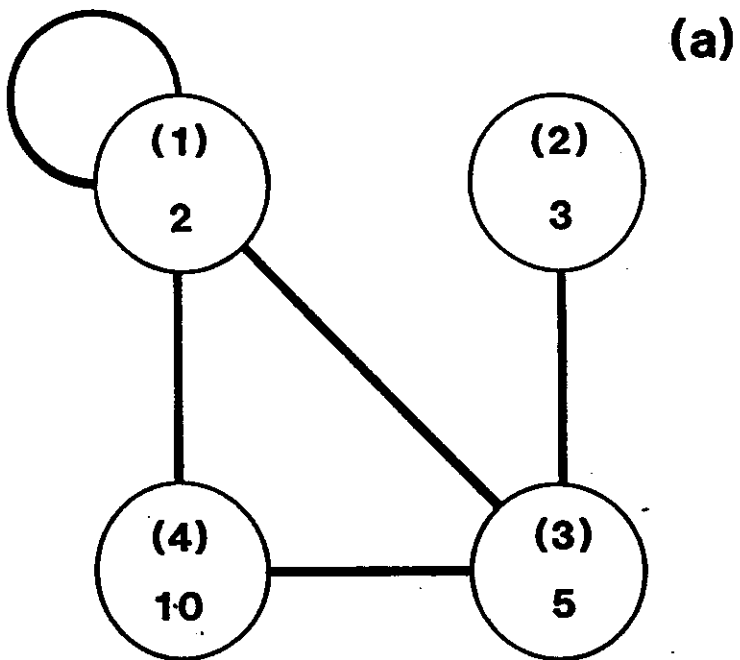
$$\underset{\sim}{X} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

(b)
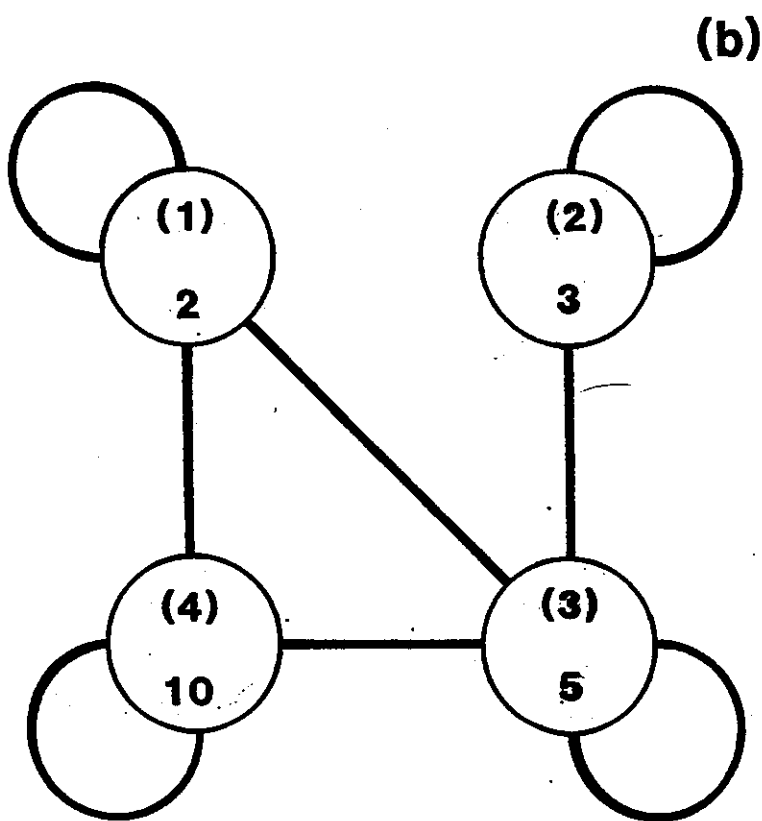
$$\underset{\sim}{X} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$
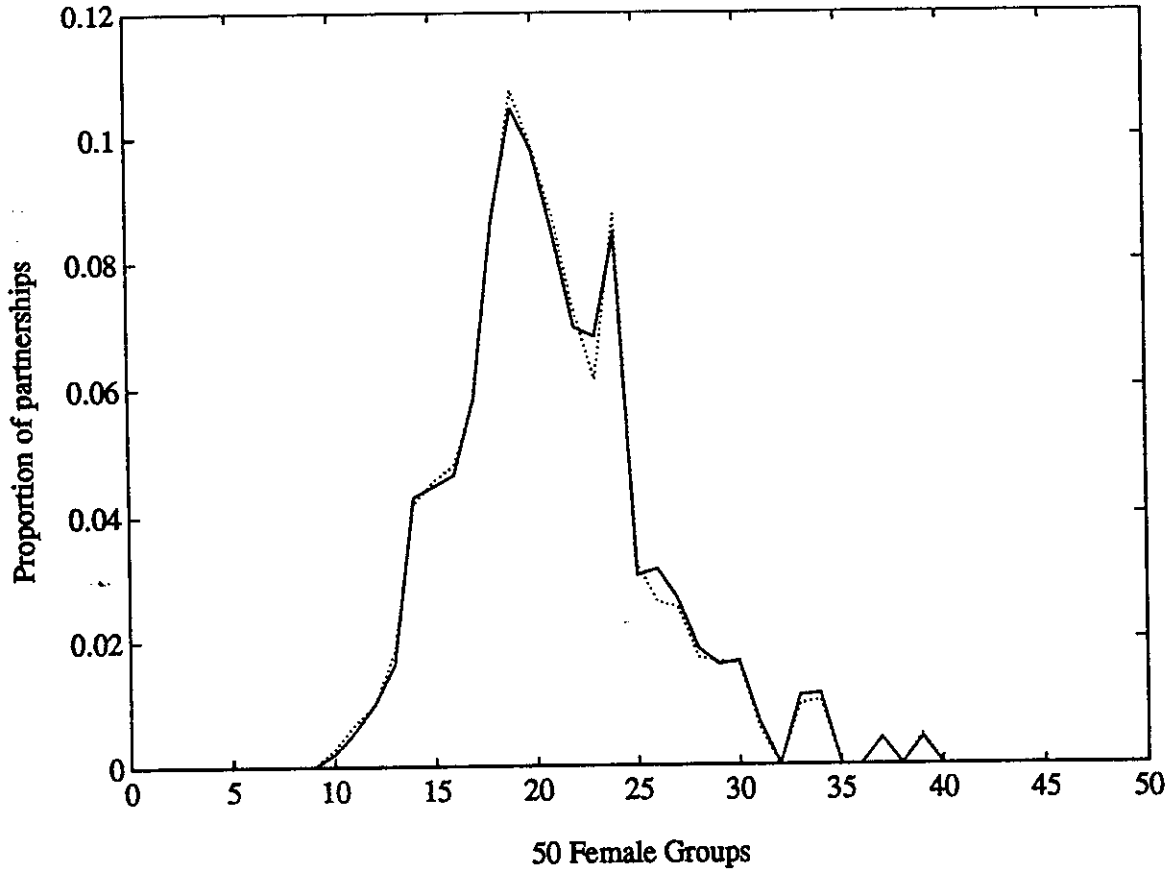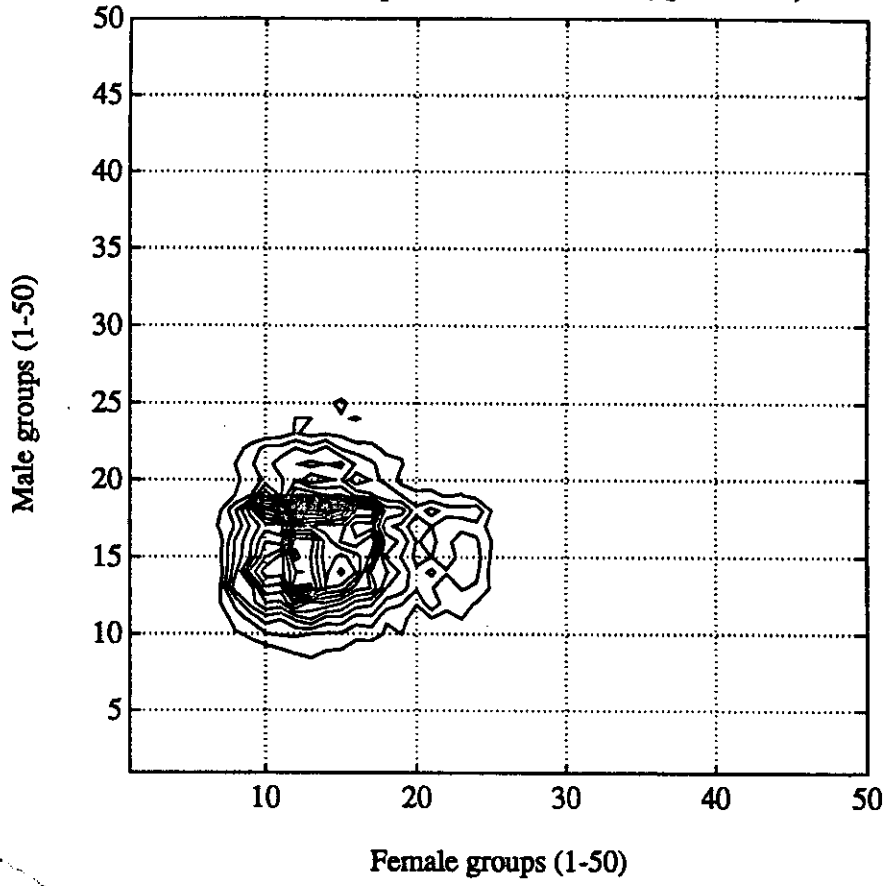
Fig 6
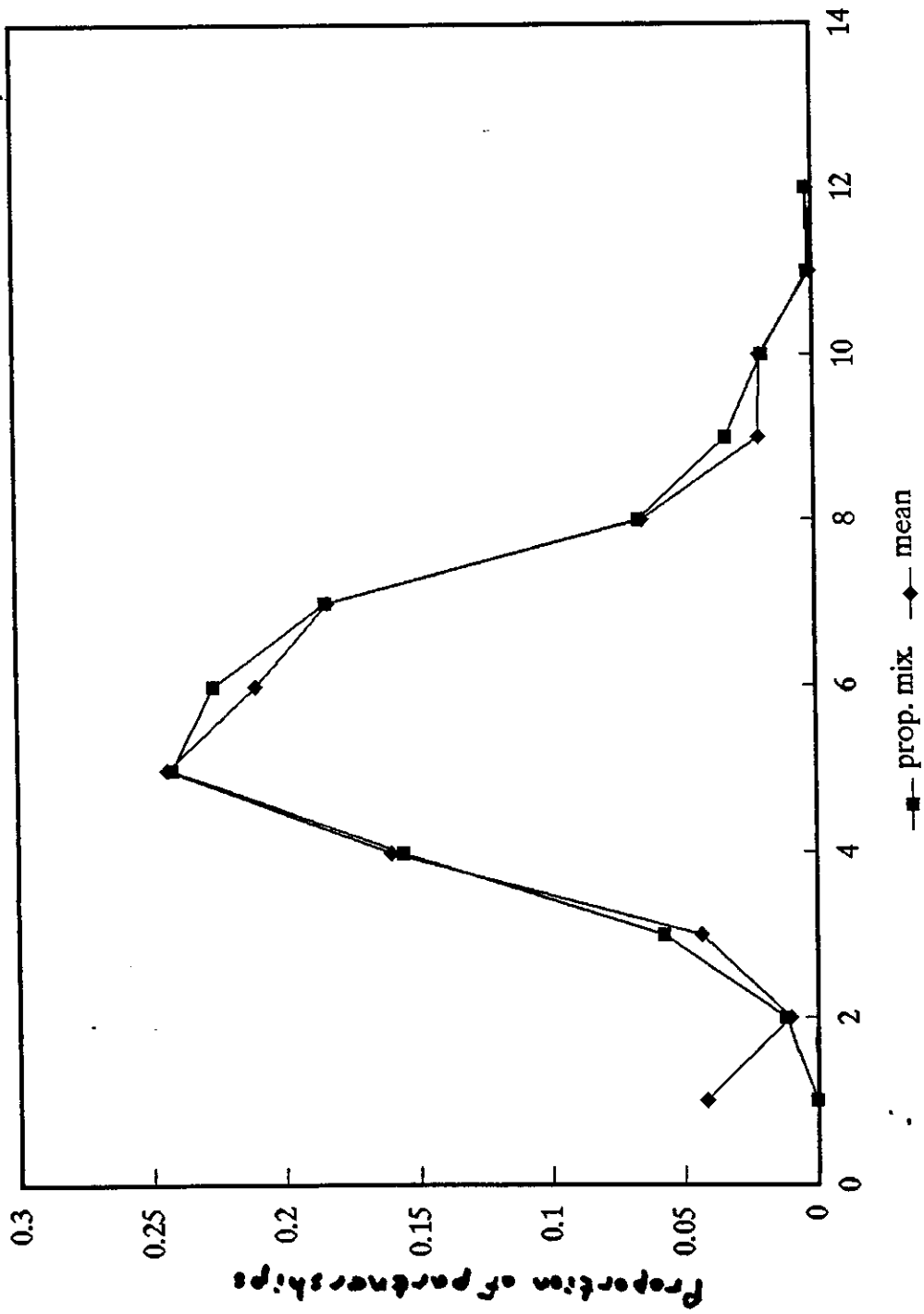
Pjbar(solid line) vs. Pjhat(dotted line) for Q = 1.0E+0

The contour of pair formation matrix (Q = 1.0E+1)

Male groups (1-50)

Female groups (1-50)

The contour of pair formation matrix (Q = 1.0E+0)

Male groups (1-50)

Female groups (1-50)

42

The contour of pair formation matrix (Q = 1.0E-7)

Female groups (1-50)

Male groups (1-50)
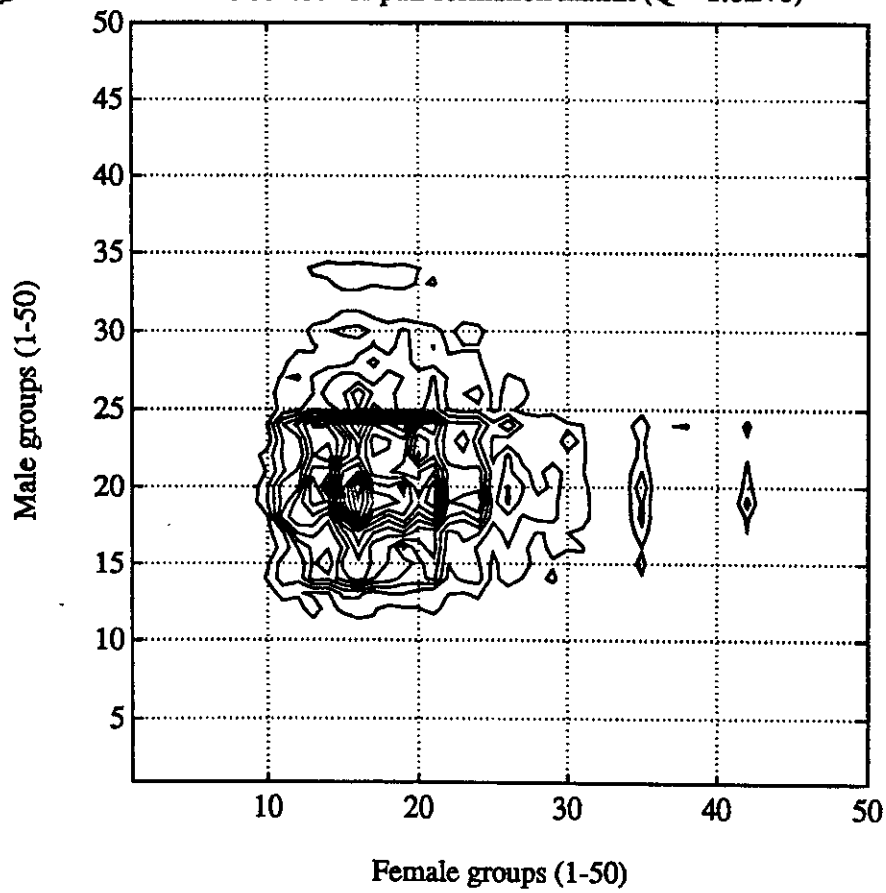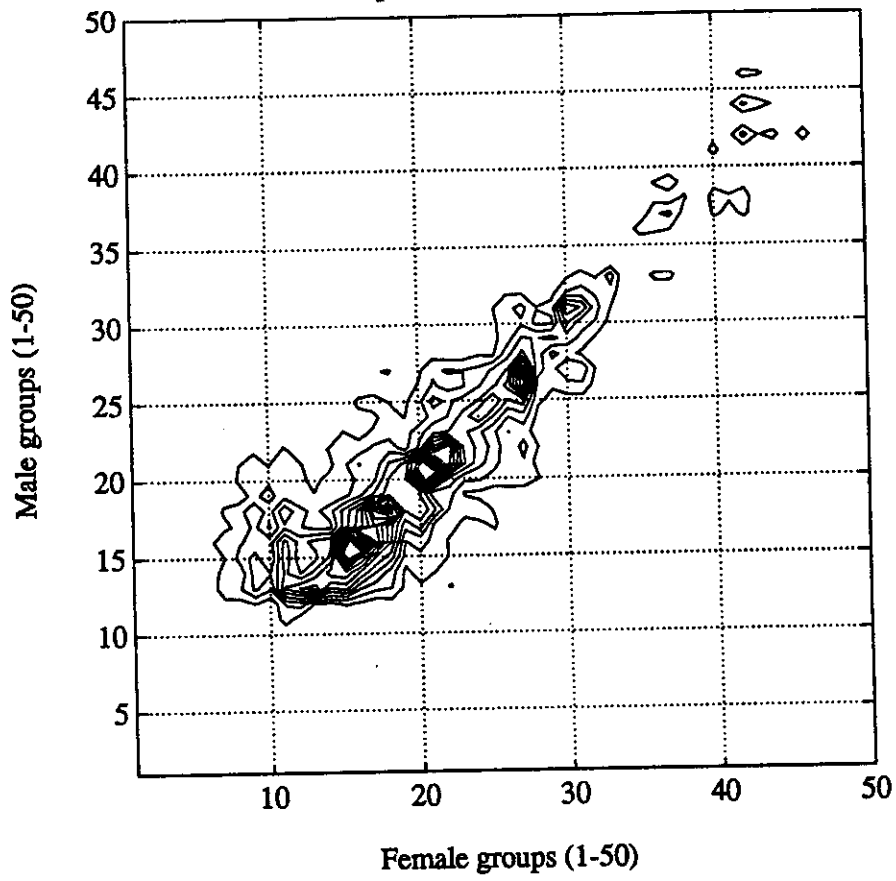
Correlation coefficient r vs. Log10(Q)

Fig 8

Fig 9

44

Fig 9

In-mixing ■   Out-mixing ◆

45