

# A Framework for Modeling Inheritance of Social Traits<sup>†</sup>

*S. Lubkin*<sup>\*1</sup>, *S.-F. Hsu Schmitz*<sup>2</sup> and *C. Castillo-Chavez*<sup>3</sup>

**Abstract.** Transmission of cultural traits behaves superficially like genetic transmission, but is substantially more complicated, since transmission is influenced by the population at large, as is disease transmission. We present a framework for modeling cultural transmission by a system of ordinary differential equations, with nonlinearities both in the transmission and in the formation of pairs. The framework is illustrated with a simple example, which can be analyzed in depth. The importance of tying models to data is emphasized, and we discuss the subtleties of collecting and interpreting the data. An example of determination of real mixing patterns is given.

Key words: pair formation, two-sex population, estimation

AMS Subject Classification: 92H10

## 1. Introduction

Mathematical models of social dynamics have been studied extensively in the context of demography and in their own right. There is an extensive literature on marriage functions that goes back to the work of Kendall (1949). Currently, there are two approaches that dominate the modeling of social dynamics. The first—classical dynamics—follows the birth and death processes within the female population. This approach is based on the work of MacKendrick (1926), Lotka (1923), and Leslie (1945). The Lotka-Volterra-MacKendrick formalism has proven extremely useful in the study of the effects of age-dependent mortality and fertility in demographic processes. An alternative approach that is gaining considerable attention consists of using models that follow the dynamics of pairs. This formalism naturally incorporates the processes of pair-formation and dissolution in addition to the usual birth and death processes. Pair formation models were introduced to study demographic

---

† This research has been partially supported by NSF grant DEB-9253570

\* To whom correspondence is to be sent.

1 Dept. of Mathematics and Statistics, Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260.

2 Biometrics Unit, 337 Warren Hall, Cornell University, Ithaca, NY 14853.

3 Biometrics Unit and Center for Applied Mathematics, 337 Warren Hall, Cornell University, Ithaca, NY 14853.

processes by Kendall (1949), Keyfitz (1949), Parlett (1972), Fredrickson (1971), McFarland (1972) and Pollard (1973); population genetics processes by Gimelfarb (1988a,b)—and many others; epidemiological processes by Dietz and Hadelar (1988), Dietz (1988), Hadelar (1989a,b), Hadelar and Ngoma (1990), Waldstätter (1989); and social dynamics and epidemiological processes by Castillo-Chavez (1989), Busenberg and Castillo-Chavez (1989, 1991), Castillo-Chavez and Busenberg (1991), Castillo-Chavez *et al.* (1991), Blythe and Castillo-Chavez (1989), and Castillo-Chavez and Blythe (1989).

The availability of an axiomatic framework (see Busenberg and Castillo-Chavez 1989, 1991 and Castillo-Chavez and Busenberg 1991) opens up the possibility for systematic study of a variety of questions in demography, epidemiology, population genetics, ecology, and social and cultural dynamics. Our objective here is to illustrate the use of our axiomatic approach for the construction of dynamic models that may prove useful in the study of the propagation or survival of social traits such as religion and language. For alternative approaches see the works of Cavalli-Sforza and Feldman (1981), and Boyd and Richerson (1985), and Baggs and Freedman (1990). In this manuscript we outline a unified modeling approach to social and cultural dynamics that incorporates the dynamics of “couples”. Section 2 introduces classical modeling approaches for pair-dynamics as described in Kendall (1949), Fredrickson (1971), McFarland (1972), Dietz and Hadelar (1988) as well as the more recent approach to pair-formation dynamics introduced by Blythe and Castillo-Chavez (1989), Castillo-Chavez and Blythe (1989), Busenberg and Castillo-Chavez (1989, 1991), Castillo-Chavez and Busenberg (1991), Castillo-Chavez *et al.* (1991). Section 3 introduces our formalism in a demographic setting. The flexibility of this formalism is illustrated in some simple cultural settings. The objective here is not to provide very detailed models but rather to show how one constructs models to address specific questions. Section 4 presents an analysis of one of these basic models. Section 5 indicates possible approaches for connecting these models with data.

## 2. Classical and Modern Approaches

In 1972, Parlett asked “Can there be a marriage function?” (Parlett 1972). What he asked in mathematical parlance is whether or not it is possible to have a satisfactory mathematical description of heterosexual pair-formation. He addressed the issue of whether there is a large class of functions or functionals which can be used to characterize mathematically valid and biologically relevant descriptions of this process. The first answer—known to us—is provided by what we refer to as the classical demographic pair-formation model. It was introduced, modified, developed, and analyzed by Kendall (1949), Fredrickson (1971), Dietz and Hadelar (1988), and Dietz (1988). Their approach is based on the use of a nonlinear function  $\psi$  to model the process (rate) of pair formation. For the situation with one class each of unpaired females  $f(t)$  and males  $m(t)$ , the mixing/pair formation function

$\psi$  for this heterosexually-active mixing population is assumed to satisfy the Fredrickson/McFarland (1971, 1972) properties at time  $t$  (we suppress the time argument for simplicity):

$$(a) \psi(0, f) = \psi(m, 0) = 0,$$

that is, in the absence of either males or females there will be no heterosexual pair formation;

$$(b) \psi(\alpha m, \alpha f) = \alpha \psi(m, f) \text{ for all } \alpha, m, f \geq 0,$$

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

$$(c) \psi(m + u, f + v) \geq \psi(m, f) \text{ for all } u, v, m, f \geq 0,$$

that is, observed increases in the number of males and/or females do not decrease the rate of pair formation.

Condition (b) implies that most reasonable mixing functions are of the form

$$\psi(m, f) = mg\left(\frac{f}{m}\right) = fh\left(\frac{m}{f}\right), \quad (1)$$

where  $h(\cdot)$  and  $g(\cdot)$  are arbitrary differentiable functions of one variable.

In demography and epidemiology, researchers have employed a variety of pair-formation or mixing functions including

$$\psi(m, f) = c \min\{m, f\},$$

$$\psi(m, f) = c\sqrt{mf},$$

$$\psi(m, f) = 2c\frac{mf}{m+f},$$

where  $c$ , an arbitrary positive constant, denotes the rate of pair formation.

The mixing, contact, or pair-formation function describes the proportion of sexual partnerships, marriages, etc. between males and females per unit time. If one wishes to generalize to situations in which age or social or cultural factors become an important mixing criterion, then one can easily extend the above approach to include these internal variables.

In this paper, we will consider cultural traits to be an important factor in the pair-formation process, which in turn becomes important in the transmission of cultural traits. We will use the mixing axioms of Castillo-Chavez and Busenberg (1991), as described below.

### 3. Formulation of the General Model

Consider a two-sex population divided into  $N$  groups. The groups could represent different native languages, religions, socio-economic groups, or even geographic characteristics. For simplicity, we will assume that the number of groups of males is the same as the number of groups of females, a 50:50 sex ratio, identical mortality

rate for all individuals, and equal birth rates for males and females. These assumptions can be easily relaxed but such generalizations will be avoided in this paper, as we wish to get our approach across as clearly as possible.

Let

$N$  = number of male groups = number of female groups

$i, j = 1, 2, \dots, N$

$m_i$  = number of single males in group  $i$  at time  $t$

$f_j$  = number of single females in group  $j$  at time  $t$

$Q_{ij}$  = number of pairs, males of type  $i$  and females of type  $j$  at time  $t$

$M_i = \sum_j Q_{ij}$  = number of paired males of type  $i$  at time  $t$

$F_j = \sum_i Q_{ij}$  = number of paired females of type  $j$  at time  $t$

$X_j^{lk}$  = probability (or proportion) that an offspring of (male- $l$ , female- $k$ ) pair is of type  $j$

and consequently  $\sum_{j=1}^N X_j^{lk} = 1$

$r_{lk} \equiv$  reproductive rate of  $l$ - $k$  pair

$\mu \equiv$  mortality rate

$b_j \equiv$  per capita pair formation rate for females of type  $j$

$c_i \equiv$  per capita pair formation rate for males of type  $i$

$p_{ij} \equiv$  probability that a male of type  $i$  pairs with a female of type  $j$  given that he pairs

$q_{ji} \equiv$  probability that a female of type  $j$  pairs with a male of type  $i$  given that she pairs

$c_i p_{ij} \equiv$  per capita rate of loss of single  $i$  males to  $(i, j)$  partnerships

$b_j q_{ji} \equiv$  per capita rate of loss of single  $j$  females to  $(i, j)$  partnerships

$\sigma_{ij} \equiv$  per capita separation rate of  $(i, j)$  pairs

We will use the fact that

$$\sum_{j=1}^N p_{ij} = \sum_{i=1}^N q_{ji} = 1; \quad p_{ij} \geq 0, \quad q_{ji} \geq 0.$$

The recruitment rate of single females or males of type  $j$  is given by

$$\sum_{l=1}^N \sum_{k=1}^N r_{lk} X_j^{lk} Q_{lk} = \Lambda_j.$$

Following Castillo-Chavez and Busenberg (1991), we use an axiomatic framework to describe the probabilities associated with pair-formation, which form the basis of our models of the dynamics of cultural traits. Specifically, the set of mixing probabilities  $\{p_{ij}(t) \text{ and } q_{ji}(t) : i = 1, \dots, N \text{ and } j = 1, \dots, N\}$  establishes the mixing/pair formation among individuals of a heterosexually-active population if they satisfy the following (postulated) set of properties:

**Definition.**  $(p_{ij}(t), q_{ji}(t))$  is called a mixing/pair-formation matrix if and only if it satisfies the following properties at all times:

$$(A1) \ 0 \leq p_{ij} \leq 1, \quad 0 \leq q_{ji} \leq 1,$$

$$(A2) \ \sum_{j=1}^N p_{ij} = 1 = \sum_{i=1}^N q_{ji},$$

$$(A3) \ c_i m_i p_{ij} = b_j f_j q_{ji}, \quad i = 1, \dots, N, \quad j = 1, \dots, N.$$

$$(A4) \ \text{If for some } i, 1 \leq i \leq N \text{ and/or some } j, 1 \leq j \leq N, \quad c_i b_j m_i f_j = 0, \text{ then we define } p_{ij} \equiv q_{ji} \equiv 0.$$

Property (A3) can be interpreted as a conservation of partnerships law or a group reversibility property (applied to rates), while (A4) asserts, the obvious, that is, that the mixing of “non-existing” or non-sexually active subpopulations cannot be arbitrarily defined. A useful class of solutions is given by Ross solutions. These solutions correspond to proportionate mixing when there are two clearly distinct sets of individuals who do not mix among themselves. Ross solutions naturally arise if we search for separable solutions (see Castillo-Chavez and Busenberg 1991).

**Definition.** A two-sex mixing/pair-formation function is called separable if and only if

$$p_{ij} = p_i p^j \quad \text{and} \quad q_{ji} = q_j q^i.$$

**Theorem 1.** The only separable pair-formation function is the Ross solution given by  $(\bar{p}^j, \bar{q}^i)$  where

$$\bar{p}^j = \frac{b_j f_j}{\sum_{i=1}^N c_i m_i}, \quad \bar{q}^i = \frac{c_i m_i}{\sum_{j=1}^N b_j f_j}; \quad j = 1, \dots, N \quad \text{and} \quad i = 1, \dots, N. \quad (2)$$

All solutions to the above axioms can be expressed as multiplicative perturbations of Ross solutions. This result is expressed in the following theorem.

**Theorem 2.** Let  $\{\phi_{ij}^m\}$  and  $\{\phi_{ji}^f\}$  be two nonnegative matrices. Let  $\ell_i^m \equiv \sum_{k=1}^N \bar{p}^k \phi_{ik}^m$  and  $\ell_j^f \equiv \sum_{k=1}^N \bar{q}^k \phi_{jk}^f$ , where  $\{(\bar{p}^j, \bar{q}^i) \mid j = 1, \dots, N \text{ and } i = 1, \dots, N\}$  denote the set of Ross solutions. Let  $R_i^m \equiv 1 - \ell_i^m$ ,  $i = 1, \dots, N$  and  $R_j^f \equiv 1 - \ell_j^f$ ,  $j = 1, \dots, N$ , and assume that  $\phi_{ij}^m$  and  $\phi_{ji}^f$  are chosen in such a way that  $R_i^m$  and  $R_j^f$  remain nonnegative for all time. Further assume that

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

and

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

Then all solutions to axioms (A1)–(A4) are given (formally) by the following multiplicative perturbations to the separable mixing solution  $(\bar{p}^j, \bar{q}^i)$ :

$$p_{ij} = \bar{p}^j \left[ \frac{R_j^f R_i^m}{\sum_{k=1}^N \bar{p}^k R_k^f} + \phi_{ij}^m \right], \quad (3a)$$

$$q_{ji} = \bar{q}^i \left[ \frac{R_i^m R_j^f}{\sum_{k=1}^N \bar{q}^k R_k^m} + \phi_{ji}^f \right], \quad (3b)$$

$$i = 1, \dots, N; \quad j = 1, \dots, N.$$

(The formal proof of this result can be found in Castillo-Chavez and Busenberg 1991.)

With the above mixing framework, we can now write the general model:

$$\frac{dm_i}{dt} = \Lambda_i - (\mu + c_i)m_i + \sum_{j=1}^N (\mu + \sigma_{ij})Q_{ij}, \quad (4a)$$

$$\frac{df_j}{dt} = \Lambda_j - (\mu + b_j)f_j + \sum_{i=1}^N (\mu + \sigma_{ij})Q_{ij}, \quad (4b)$$

$$\frac{dQ_{ij}}{dt} = c_i m_i p_{ij} - (2\mu + \sigma_{ij})Q_{ij} = b_j f_j q_{ji} - (2\mu + \sigma_{ij})Q_{ij}. \quad (4c)$$

Where  $\Lambda_i, \Lambda_j, p_{ij}, q_{ji}$  may be complicated nonlinear functions of the state variables and/or time and where we have made use of the conservation of rates property (A3).

## 4. Maternal Determination Model

As an example, consider the following model. Let the mortality ( $\mu$ ) and separation rates ( $\sigma$ ) be the same for all groups, and let the type of offspring be determined solely by the type of its mother, that is,

$$X_j^{lk} = \delta(k, j) \equiv \begin{cases} 1 & \text{if } k = j; \\ 0 & \text{if } k \neq j. \end{cases}$$

Furthermore, let the recruitment rate be directly proportional to the number of females of each type with the birth rate for females of type  $i$  denoted by  $r_i$ . Then the general model reduces to the following set of equations

$$\frac{dm_i}{dt} = r_i F_i - (\mu + c_i)m_i + (\mu + \sigma)M_i, \quad (5a)$$

$$\frac{df_j}{dt} = r_j F_j - (\mu + b_j)f_j + (\mu + \sigma)F_j, \quad (5b)$$

$$\frac{dQ_{ij}}{dt} = c_i m_i p_{ij} - (2\mu + \sigma)Q_{ij} \quad (5c)$$

$$= b_j f_j q_{ji} - (2\mu + \sigma)Q_{ij}, \quad (5d)$$

where, recall,  $M_i = \sum_j Q_{ij}$  and  $F_j = \sum_i Q_{ij}$ .

Summing equation (5c) above over  $j$  and (5d) over  $i$  we obtain the following model for the aggregated dynamics:

$$\frac{dm_i}{dt} = r_i F_i - (\mu + c_i) m_i + (\mu + \sigma) M_i, \quad (6a)$$

$$\frac{df_j}{dt} = r_j F_j - (\mu + b_j) f_j + (\mu + \sigma) F_j, \quad (6b)$$

$$\frac{dM_i}{dt} = -(2\mu + \sigma) M_i + c_i m_i, \quad (6c)$$

$$\frac{dF_j}{dt} = -(2\mu + \sigma) F_j + b_j f_j, \quad (6d)$$

since, by definition,  $\sum_j p_{ij} = \sum_i q_{ji} = 1$ .

System (6) is a linear system, so its only equilibrium is zero, and all its solutions are exponential. We can write (6) in matrix form:

$$\dot{X} = AX \quad (7)$$

where  $X = (m_1, \dots, m_N, f_1, \dots, f_N, M_1, \dots, M_N, F_1, \dots, F_N)^T$  and where  $A$ , a  $(4N \times 4N)$  matrix, can be written in block form as

$$A = \begin{bmatrix} B & 0 & C & D \\ 0 & E & 0 & F \\ G & 0 & H & 0 \\ 0 & J & 0 & K \end{bmatrix} \quad (8)$$

where the  $(N \times N)$  submatrices  $B, C, D, E, F, H, J, K$  are given explicitly by

$$B = \text{diag}(-\mu - c_i),$$

$$C = \text{diag}(\mu + \sigma),$$

$$D = \text{diag}(r_i),$$

$$E = \text{diag}(-\mu - b_j),$$

$$F = \text{diag}(\mu + \sigma + r_j),$$

$$G = \text{diag}(c_i),$$

$$H = \text{diag}(-(2\mu + \sigma)),$$

$$J = \text{diag}(b_j),$$

$$K = \text{diag}(-(2\mu + \sigma)) = H.$$

Consequently,  $A$  is a block matrix where each block is diagonal.  $A$  is easily triangulable by row operations:

$$\tilde{A} = \begin{bmatrix} B & 0 & C & D \\ 0 & E & 0 & F \\ 0 & 0 & H - \left[\frac{G}{B}\right] C & -\left[\frac{G}{B}\right] D \\ 0 & 0 & 0 & K - \left[\frac{J}{E}\right] F \end{bmatrix} \quad (9)$$

where

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{ij} = \begin{cases} \frac{\alpha_{ij}}{\beta_{ij}}, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

*i.e.*, scalar division of nonzero matrix elements. The eigenvalues of  $A$  must be the same as those of  $\tilde{A}$ , which are immediately seen to be

$$\lambda = -\mu - c_i, \quad -\mu - b_j, \quad \frac{c_i}{\mu + c_i}(\mu + \sigma) - (2\mu + \sigma), \quad \frac{b_j}{\mu + b_j}(\mu + \sigma + r_j) - (2\mu + \sigma), \quad (10)$$

$i, j = 1, \dots, N$ . The first  $2N$  of these are always negative. The next  $N$  of these are seen to be always negative as well:

$$\begin{aligned} \lambda &= \frac{c_i}{\mu + c_i}(\mu + \sigma) - (2\mu + \sigma) \\ &= \frac{-\mu(2\mu + \sigma + c_i)}{\mu + c_i} < 0. \end{aligned} \quad (11)$$

But the last  $N$  eigenvalues may take on positive values if  $r_j$  is sufficiently large. Thus the stability of the zero equilibrium will depend on the birth rates  $r_j$ .

A simple estimate of the range of  $\lambda$  is obtained by looking at  $\lambda$  as the pair formation rate  $b_j$  gets arbitrarily large:  $\lambda \sim (\mu + \sigma + r_j) - (2\mu + \sigma) = r_j - \mu$  so that  $r_j > \mu$  implies  $\lambda > 0$  in the limit as  $b_j \rightarrow \infty$ . That is to say, if pair formation is instantaneous, separation is rendered demographically meaningless, and growth or decay are determined by the gross difference between the birth and mortality rates.

An alternative estimate for the spectrum of  $A$  is obtained by observing that the first  $3N$  columns of  $A$  are strictly diagonally dominant, and that the last  $N$  columns are strictly diagonally dominant if  $\mu + \sigma + 2r_i < 2\mu + \sigma \quad \forall i$ . That is, a necessary and sufficient condition for the strict diagonal dominance of  $A$  is that  $\mu > 2 \max(r_i)$ . Observing further that all the diagonal elements of  $A$  are negative, and applying the Levy-Desplanques Theorem (see R. Horn and C. Johnson, 1985) we see that a sufficient condition for all eigenvalues of  $A$  to have negative real part is that  $\mu > 2 \max(r_i)$ . This, again, is an intuitive result but does not require the artificial device of instantaneous pair formation. In summary, for the aggregated linear model (6), the zero solution is stable if  $\mu > 2 \max(r_i)$ , *i.e.*, mortality outpaces reproduction.

## 5. Connections to data

A problem of considerable theoretical importance lying at the interface of social and cultural dynamics, demography, and epidemiology is determining and modeling who is interacting or mixing with whom. Unravelling the social/cultural structure of this network of interactions in a real setting is complicated by a variety of factors, including the problem of estimating the effective population and/or group size of the network under

consideration. Determining the effective size, that is, the number of individuals who are active members of a social/cultural network is extremely difficult because we do not know whether or not we are working with a closed social network and this fact is usually ignored in most theoretical work. We (Rubin *et al.* 1992, Castillo-Chavez *et al.* 1992) have used a mark-recapture paradigm to estimate the sizes of the populations having social (sexual) contacts with a specified population and hence susceptible to disease/cultural transmission. The need to estimate the size of the socially (sexually) active subset before estimating the size of the out-mixing population introduces extra variability into the problem. An estimator of the variance of the estimated size of the population at risk that accounts for this extra variability and an expression for the bias of such an estimator have been derived (see above references). We illustrate our results with data on “dating” collected from a population of university undergraduates, and make use of our axiomatic modeling approach for mixing/pair formation to compute specific mixing matrices (the details will be published elsewhere).

A survey of social and sexual patterns among college students (see Crawford *et al.* 1990) reveals that about 50% of the sexual relationships of students involve partners who are not college students. We will refer to the class of non-students that interact socially or sexually with our college population as the *other* or *unknown* class (see Castillo-Chavez *et al.* 1992). If the number of social and/or sexual contacts with the *other* population is significant, then mixing/contact matrices involving only the population of interest would not be sufficient to study the dynamics of cultural, sociological, or sexually-transmitted traits.

However, the explicit computation of mixing matrices of a non-closed population involves the estimation of many parameters, including the size of the active *other* population and the proportions of the relationships (social/sexual contacts) that individuals in the known population had with the *other* group. The former can be estimated by using mark-recapture methodology, as shown by Rubin *et al.* (1992), while the latter estimation problem can be reduced to the estimation of a single parameter through the systematic application of the two-sex mixing framework (see Castillo-Chavez *et al.* 1992 and Castillo-Chavez and Busenberg, 1991). In the remainder of this section we summarize this approach using our data on heterosexually intimate relationships of college students, which are more general than sexual relationships, *e.g.*, dating.

The data is from a survey of university undergraduates, which was conducted by Crawford *et al.* (1990). Only those responses to some questions related to heterosexual dating are presented here. Subjects (respondents) are categorized into four classes: 1 (freshman), 2 (sophomore), 3 (junior), and 4 (senior). Students in these four classes are members of the known college undergraduate population. Their partners are categorized into the above four classes plus an additional fifth class (the *other*) to take into account “external” contacts/partners. We use the superscripts  $m$  and  $f$  to indicate the male and female populations, respectively (note that  $m$  and  $f$  have different meaning when they are not superscripts), and the subscripts  $i$  and  $j$  to indicate their respective class. The exact population sizes of the four known classes, here denoted by  $R_i^m$  and  $R_j^f$ , are available from

Table 1. Population sizes and sample sizes for males/females

Class	Popn. Size	Sample Size	Act. Sample Size	Act. Prop.	Act. Popn. Size
$i/j$	$R$	$S$	$A$	$A \div S$	$T$
1	1673/1278	79/ 66	56/ 44	0.709/0.667	1186/ 852
2	1589/1308	60/ 57	38/ 45	0.633/0.789	1006/1033
3	1591/1277	38/ 51	29/ 39	0.763/0.765	1214/ 977
4	1686/1348	73/ 65	51/ 47	0.699/0.725	1178/ 975
Total	6539/5211	250/239	174/175	0.696/0.732	4584/3837

Table 2. Dating partnerships distribution  $U_{ij}^m$ , mixing proportions ( $p_{ij}$ ) and average number of partners  $c_i$  for male respondents

Class	Female		Partner	Class	$j$	Total	Average	
$i$	1	2	3	4	$X_i^m$	5	$c_i$	
1	123	26	15	4	168	50	218	3.89
	(0.564)	(0.119)	(0.069)	(0.018)		(0.229)		
2	23	53	20	7	103	25	128	3.37
	(0.180)	(0.414)	(0.156)	(0.055)		(0.195)		
3	11	19	27	15	72	19	91	3.14
	(0.121)	(0.209)	(0.297)	(0.165)		(0.209)		
4	11	11	27	53	102	39	141	2.76
	(0.078)	(0.078)	(0.191)	(0.376)		(0.277)		
Total	168	109	89	79	445	133	578	3.32
	(0.291)	(0.189)	(0.154)	(0.137)	$X_+^m$	(0.230)	$Y_+^m$	

Table 3. Dating partnerships distribution  $U_{ji}^f$ , mixing proportions ( $q_{ji}$ ) and average number of partners  $b_j$  for female respondents

Class	Male		Partner	Class	$i$	Total	Average	
$j$	1	2	3	4	$X_j^f$	5	$b_j$	
1	58	17	6	9	90	24	114	2.59
	(0.509)	(0.149)	(0.053)	(0.079)		(0.211)		
2	9	45	27	15	96	33	129	2.87
	(0.070)	(0.349)	(0.209)	(0.116)		(0.256)		
3	2	10	41	22	75	22	97	2.49
	(0.021)	(0.103)	(0.423)	(0.227)		(0.227)		
4	2	6	20	45	73	53	126	2.66
	(0.016)	(0.048)	(0.159)	(0.357)		(0.421)		
Total	71	78	94	91	334	132	466	2.66
	(0.152)	(0.167)	(0.202)	(0.195)	$X_+^f$	(0.283)	$Y_+^f$	

the university registrar office. The sample sizes, here denoted by  $S_i^m$  and  $S_j^f$ , provide a count of the number of survey respondents, while the active sample sizes, denoted by  $A_i^m$  and  $A_j^f$ , provide a count of the number of respondents who were active in dating during the two-month period (our unit of time) prior to the survey. For each subject class, the total number of partners (distinct individuals with whom the respondent experienced “dating” activity during a specific two-month period) is denoted by  $Y_i^m$  and  $Y_j^f$  while the class distribution of dating activity among each of the five partner classes (as self-reported) is denoted by  $U_{ij}^m$  and  $U_{ji}^f$ . The partnerships involving partners exclusively from the first four classes are denoted by  $X_i^m$  and  $X_j^f$ . They are calculated by summing the partnerships of the respondents with classes 1 through 4. Dividing  $U_{ij}^m$  and  $U_{ji}^f$  by their corresponding  $Y_i^m$  and  $Y_j^f$  results in the mixing proportions, here denoted by  $p_{ij}$  for males and  $q_{ji}$  for females. The average number of partners per unit time for active subjects, here denoted by  $c_i$  and  $b_j$ , are estimated by dividing  $Y_i^m$  and  $Y_j^f$  by  $A_i^m$  and  $A_j^f$ , respectively. Table 1 summarizes these statistics as well as the estimated active population sizes (denoted by  $T_i^m$  and  $T_j^f$ ) in each of the four subject classes. Estimated active population sizes are calculated by multiplying the population size by the active proportion in the sample. Among respondents, about 70% were active in dating with the overall active proportion for females (73.2%) slightly higher than that for males (69.6%). The partnership distribution, mixing proportions, and average number of partners are presented in Tables 2 and 3. The row total of the mixing proportions may not be exactly equal to 1 because of rounding errors.

For active males (females), the overall proportion of partnerships with class 5 (*other*) is 23.0% (28.3%) and the overall average number of partners is 3.32 (2.66) per unit time (2 months). If the data is to be consistent, then females of known classes must have more distinct partners from the *other* class than males of known classes. These two mixing matrices are plotted in Figures 1 and 2.

From these figures we conclude that qualitatively there is a like-with-like mixing pattern among male and female individuals of the known population, that is, freshmen prefer freshmen, sophomores prefer sophomores, etc. In addition, we observe that males mix more with females of the same or lower college classes, while females mix more with males of the same or higher classes. Since the age of our respondents highly correlates with the class that they belong to, we conclude that males prefer to date younger females and females prefer older males. Finally, we observe that the proportion of dating activity with the *other* population is substantial. A more detailed characterization of this population requires further surveys.

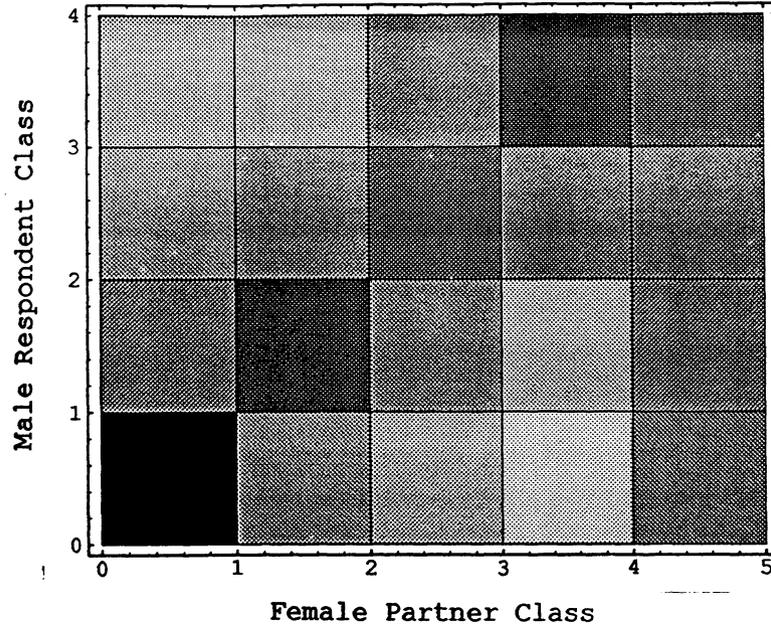


Figure 1. Dating pattern of male respondents

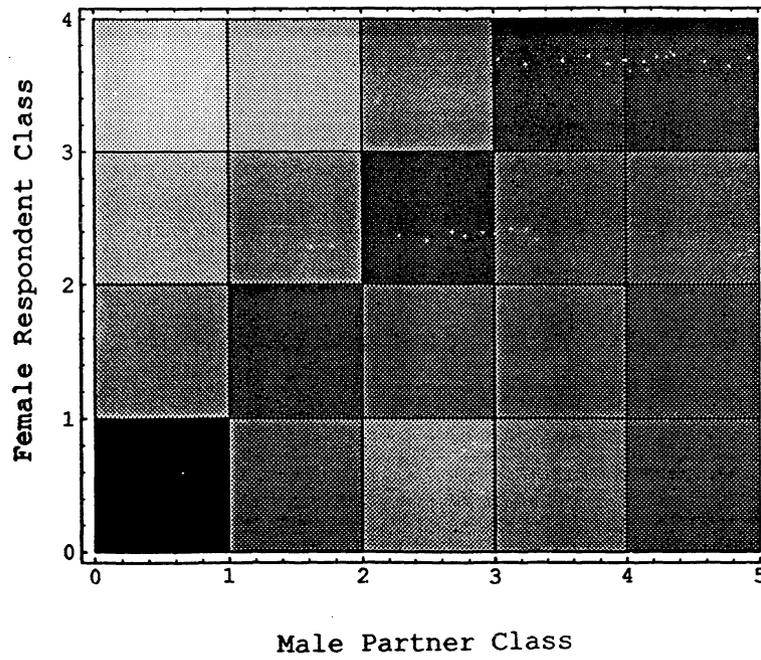


Figure 2. Dating pattern of female respondents

## 6. Conclusions

Previous approaches to modeling cultural transmission have either been very specific, as in the bilingual competition model of Baggs and Freedman (1990), or have been general, but confined by a rigid framework. For example, Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985) both construct an elaborate framework based on cultural transmission once per generation, *e.g.*, at birth. This has the disadvantages of excluding cultural transmission during an individual's lifetime (*e.g.*, religious conversion) and of requiring that generations be of equal duration and generated simultaneously. The discrete-generation approach is not necessarily any easier to analyze than our differential equation approach, and it may introduce behaviors (*e.g.*, periodicity) more appropriate to organisms which do have discrete, simultaneous generations, such as insects, than to humans.

The approach we follow here is built on first principles (Axioms A1–A4) and a growing methodology tied in to data (see Castillo-Chavez *et al.* 1991, Castillo-Chavez *et al.* 1992, and Rubin *et al.* 1992). This methodology begins with survey data and is based on the work of Bailey (1951) and Seber (1982), and more recently on the work on empirical Bayes estimation outlined in Castillo-Chavez *et al.* (1991).

The problems of cultural transmission are complex, and the data is minimal. We plan to continue this work in several directions. Nonlinearities arise in our work at two fronts: at the recruitment level and at the transmission level. In Section 4, we analyzed a simple aggregated model. More detailed analysis will be carried out first using simple mixing structures such as the one provided by Ross solutions, though most realistic models of cultural transmission will require like-with-like mixing. For example, a study by the Council of Jewish Federations (Steinfels, 1992) reports that the religion of the spouses of American Jewish subjects was Jewish in 89% of marriages before 1965, and in 43% of marriages after 1985, although the proportion of Jews in the US population was only 2–3% during that period. Thus any serious modeling attempt must include like-with-like preferential mixing, where it is applicable. We also plan to study a variety of recruitment functions such as

- (i) “Melting pot”: (assuming 50/50 birth rates for males/females)

$$\Lambda_i = \sum_l \sum_k r_{lk} Q_{lk} \delta(k, i),$$

- (ii) Biparental determination:

$$\Lambda_i = \sum_l \sum_k r_{lk} Q_{lk} [\gamma \delta(k, i) + (1 - \gamma) \delta(l, i)],$$

where  $\gamma$  = maternal influence factor ( $0 \leq \gamma \leq 1$ ). Note this model is not much more difficult to analyze than the maternal determination model, since it has only one more linear term.

(iii) Biparental determination influenced by an outside population:

$$\Lambda_i = \sum_l \sum_k r_{lk} Q_{lk} \left\{ [\gamma\delta(k, i) + (1 - \gamma)\delta(l, i)]\beta + (1 - \beta) \frac{m_i + f_i + M_i + F_i}{\sum_h (m_h + f_h + M_h + F_h)} \right\},$$

where  $\beta$  is the proportion of parental influence, and  $(1 - \beta)$  is the proportion of outside-population influence. If  $\beta$  is small, minority extinction occurs except when there is a high minority reproductive rate.

Of course, age-structure provides the most appropriate way of incorporating the process of recruitment in our framework. Initial work in this direction has begun and we plan to report on some preliminary results soon.

## References

- [1] I. Baggs, H.I. Freedman (1990), A mathematical model for the dynamics of interactions between a unilingual and a bilingual population: persistence vs. extinction. *J. Math. Sociol.*, 16, pp. 51–75.
- [2] N.T.J. Bailey (1951), On estimating the size of mobile populations from recapture data. *Biometrika*, 38, pp. 293–306.
- [3] S.P. Blythe, C. Castillo-Chavez (1989), Like-with-like preference and sexual mixing models. *Math. Biosci.*, 96, pp. 221–238.
- [4] R. Boyd, P.J. Richerson (1985), *Culture and the Evolutionary Process*. The University of Chicago Press, Chicago.
- [5] S. Busenberg, C. Castillo-Chavez (1989), Interaction, pair formation and force of infection terms in sexually transmitted diseases. *Lect. Notes Biomath.*, 83, pp. 289–300.
- [6] S. Busenberg, C. Castillo-Chavez (1991), A general solution of the problem of mixing of subpopulations, and, its application to risk- and age- structured epidemic models. *IMA J. Math. Appl. Med. Biol.*, 8, pp. 1–29.
- [7] C. Castillo-Chavez (1989), Review of recent models of HIV/AIDS transmission. In: *Applied Mathematical Ecology* (S.A. Levin, T.G. Hallam, L.J. Gross, eds.), *Biomathematics* 18, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, pp. 253–262.
- [8] C. Castillo-Chavez, S.P. Blythe (1989), Mixing framework for social/sexual behavior. *Lect. Notes Biomath.*, 83, pp. 275–288.
- [9] C. Castillo-Chavez, S. Busenberg (1991), On the solution of the two-sex problem. In: *Proceedings of the Interantional Conference on Differential Equations and Applications to Biology and Population Dynamics* (S. Busenberg, M. Martelli, eds.), *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, Barcelona, Budapest, pp. 80–98.

- [10] C. Castillo-Chavez, S. Busenberg, K. Gerow (1991), Pair formation in structured populations. In: *Differential Equations with Applications in Biology, Physics and Engineering* (J. Goldstein, F. Kappel, W. Schappacher, eds.), Marcel Dekker, New York, pp. 47–65.
- [11] C. Castillo-Chavez, S.-F. Shyu, G. Rubin, D. Umbach (1992), On the estimation problem of mixing/pair formation matrices with applications to models for sexually-transmitted diseases. In: *AIDS Epidemiology: Methodological Issues* (N.P. Jewell, K. Dietz, V.T. Farewell, eds.), Birkhäuser, Boston, Basel, Berlin, pp. 384–402.
- [12] C.M. Crawford, S.J. Schwager, C. Castillo-Chavez (1990), A methodology for asking sensitive questions among college undergraduates. *Biometrics Unit Technical Report, BU-1105-M*, Cornell University, Ithaca, New York.
- [13] L.L. Cavalli-Sforza, M.W. Feldman (1981), *Cultural Transmission and Evolution. Monographs in Population Biology 16*. Princeton University Press, Princeton, New Jersey.
- [14] K. Dietz (1988), On the transmission dynamics of HIV. *Math. Biosci.*, 90, pp. 397–414.
- [15] K. Dietz, K.P. Hadeler (1988), Epidemiological models for sexually transmitted diseases. *J. Math. Biol.*, 26, pp. 1–25.
- [16] A.G. Fredrickson (1971), A mathematical theory of age structure in sexual populations: Random mating and monogamous marriage models. *Math. Biosci.*, 20, pp. 117–143.
- [17] A. Gimelfarb (1988a), Processes of pair formation leading to assortative mating in biological populations: encounter mating model. *Amer. Nat.*, 131(6), pp. 865–884.
- [18] A. Gimelfarb (1988b), Processes of pair formation leading to assortative mating in biological populations: dynamic interaction model. *Theor. Pop. Biol.*, 34, pp. 1–23.
- [19] S. Gupta, R.M. Anderson, R.M. May (1989), Networks of sexual contacts: implications for the pattern of spread of HIV. *AIDS*, 3, pp. 1–11.
- [20] K.P. Hadeler (1989a), Pair formation in age-structured populations. *Acta Applicandae Mathematicae*, 14, pp. 91–102.
- [21] K.P. Hadeler (1989b), Modeling AIDS in structured populations. 47th Session of the International Statistical Institute, Paris, August/September. *Conference Proc.*, C1-2.1, pp. 83–99.
- [22] K.P. Hadeler, K. Ngoma (1990), Homogeneous models for sexually transmitted diseases. *Rocky Mountain Journal of Mathematics*, 20, pp. 967–986.
- [23] R. Horn, C. Johnson (1985), *Matrix Analysis*. Cambridge Univ. Press.
- [24] D.G. Kendall (1949), Stochastic processes and population growth. *Roy. Statist. Soc., Ser.B*, 2, pp. 230–264.
- [25] N. Keyfitz (1949), The mathematics of sex and marriage. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV: Biology and Health*, pp. 89–108.

- [26] P.H. Leslie (1945), On the use of matrices in certain population mathematics. *Biometrika*, 33, pp. 183–212.
- [27] A.J. Lotka (1923), Contributions to the analysis of malaria epidemiology. *Am. J. Hyg.*, 3, Jan. Suppl.
- [28] A.G. MacKendrick (1926), Applications of mathematics to medical problems. *Proc. Edinburgh Math. Soc.*, 44, pp. 98–130.
- [29] D.D. McFarland (1972), Comparison of alternative marriage models. In: *Population Dynamics* (T.N.E. Greville, ed.), Academic Press, New York, London, pp. 89–106.
- [30] B. Parlett (1972), Can there be a marriage function? In: *Population Dynamics* (T.N.E. Greville, ed.), Academic Press, New York, London, pp. 107–135.
- [31] J.H. Pollard (1973), *Mathematical models for the growth of human populations*, Chapter 7: The two sex problem. Cambridge University Press.
- [32] G. Rubin, D. Umbach, S.-F. Shyu, C. Castillo-Chavez (1992), Application of capture-recapture methodology to estimation of size of population at risk of AIDS and/or other sexually-transmitted diseases. *Statistics in Medicine*, 11, pp. 1533–1549.
- [33] G.A.F. Seber (1982), *The Estimation of Animal Abundance and Related Parameters*. Macmillan, New York.
- [34] P. Steinfelds (1992), Debating intermarriage, and Jewish survival. *New York Times*, Oct. 16, 1992.
- [35] R. Waldstätter (1989), Pair formation in sexually transmitted diseases. *Lect. Notes Biomath.*, 83, pp. 260–274.

Sharon LUBKIN (responsible for correspondence)

Department of Mathematics and Statistics

Thackeray Hall

University of Pittsburgh

Pittsburgh, PA 15260

Tel: (412) 624-9033

Fax: (412) 624-8397

E-mail: [lubkin@next1.math.pitt.edu](mailto:lubkin@next1.math.pitt.edu)

Shu-Fang HSU SCHMITZ

Biometrics Unit

337 Warren Hall

Cornell University

Ithaca, NY 14853-7801

Tel: (607) 255-5488

Fax: (607) 255-4698

E-mail: [kwaj@cornella.cit.cornell.edu](mailto:kwaj@cornella.cit.cornell.edu)

Carlos CASTILLO-CHAVEZ

Biometrics Unit and Center for Applied Mathematics

337 Warren Hall

Cornell University

Ithaca, NY 14853-7801

Tel: (607) 255-5488

Fax: (607) 255-4698

E-mail: [carlos@carlos.cit.cornell.edu](mailto:carlos@carlos.cit.cornell.edu)