July, 2012

Models of Assortative Matching

Ted C Bergstrom, *University of California, Santa Barbara*

Models of Assortative Matching

Theodore C. Bergstrom
Economics Department
University of California Santa Barbara

July 23, 2012

1 Measures of Assortativity

Sewall Wright [8] [9] [10] defined a measure of assortative matching "for describing population structure in breeds of livestock and in natural populations." This measure has come to be known as *Wright's F-statistic.* Wright sometimes referred to the F-statistic as the *coefficient of inbreeding* in a population and, in the case of animals with known pedigree, as the *coefficient of relatedness.* Wright applied the F-statistic not only to assortative matching resulting from inbreeding of kin, but also to preferential mating patterns where mate choice is based on phenotypic similarities correlated with genotypic similarity. between them. He offered a formal definition of his F-statistic as "the correlation between homologous genes of uniting gametes under a given mating pattern,"

There is a simple connection between the F-statistic and the prevalence of "mixed pairs" in a population. In a genetics text book by Hartl and Clark, ([7], page 245) Wright's F is described as "the fractional reduction in heterozygosity relative to a random-mating population with the same allele frequencies."

In a discussion of group selection and assortative mating, Bergstrom [4] defines the *index of assortativity* of a matching process between types i and j to be the difference between the probability that one is matched with a type i *given that one is of type* i and the probability that one is matched with a type i *given that one is a type* j. In a series of papers, [1], [2], [3], Alger and Weibull apply this index of assortativity to analyze stable states in the evolution of preferences.

Theorem 1 shows that for a broad class of matching processes, the three measures, Wright's correlation, the "fractional reduction in heterozygosity" and the “index of assortativity” are all equivalent.

Definition 1. *A random matching process is an assignment such that each individual in the population is assigned exactly one partner. The probability that a randomly selected matched pair has one member of type i and one of type j is denoted π_{ij}. The conditional probability that a randomly selected individual of type j has a partner of type i is denoted $\pi(i|j)$.*

Wright’s coefficient F is the correlation between the indicator random variables for type of two individuals that are paired together. In standard genetic applications, the combinatorics of diploid sexual reproduction imply that this correlation is independent of the proportions of types in the population. However, as we will demonstrate, there are interesting combinatorial processes in which this correlation depends on the proportions of types. Thus where the proportions of the two types are p and $1-p$, we will denote Wright’s statistic as $F(p)$. The following result, whose proof is found in the appendix, shows the equivalence of Wright’s correlation, the reduction of heterozygosity, and the index of assortativity.

Proposition 1. *In a random matching process with two types, where the fractions of types 1 and 2 respectively are p and $1-p$, Wright’s correlation coefficient can be written in any of the following ways:*

$$F(p) = \frac{\pi(1|1) - p}{1-p}, \tag{1}$$

$$F(p) = 1 - \frac{\pi_{12}}{2p(1-p)}, \tag{2}$$

and as

$$F(p) = \pi(1|1) - \pi(1|2) \tag{3}$$

Equation 1 of Theorem 1 expresses $F(p)$ as the correlation coefficient of indicators for the types of two individuals who are matched. Equation 2 expresses Hartl and Clark’s “fractional reduction in heterozygosity by showing the ratio of the actual probability of a match between the two different types and the probability of such a match if matching were random is equal to $1 - F(p)$. Equation 3 shows that the “index of assortativity” is equal to Wright’s correlation coefficient.

2 Two-Pool Assortative Processes

L.L. Cavalli-Sforza and M. W. Feldman employ the F-statistic in a study of cultural evolution in a population with assortative matching ([5], page 96). They describe a matching process which can be defined as follows:

Definition 2. *A two-pool assortative matching process with uniform assortativity F is a random matching process such that with probability F, a member of the population matches from an "assortative pool", consisting only of its own type, while the complementary fraction $1-F$ matches from a "random pool" consisting of all individuals who who did not match from an assortative pool.*

Proposition 2. *Wright's F-statistic for a two-pool assortative matching process with uniform assortativity F is equal to F and is independent of the proportions p and $1-p$ in the population.*

Proof. An individual of type 1 can be matched with another type 1 in one of two ways. With probability F, the individual joins an assortative pool and is then certain to be matched with another type 1. With probability $(1-F)$, a type 1 joins the random pool. Since the fraction of type 1's in the random pool is p, the probability that a type 1 joins the random pool and is matched with a type 1 is $p(1-F)$. Therefore the conditional probability that one is matched with a type 1 given that one is of type 1 is $\pi(1|1) = F + p(1-F) = (1-p)F + p$. Substituting this expression for $\pi(1|1)$ in Equation 1 of Theorem 1, we have

$$\begin{aligned} F(p) &= \frac{(1-p)F + p - p}{1-p} \\ &= F. \end{aligned} \tag{4}$$

□

Since according to Proposition 1, the index of assortativity of a matching process is equal to Wright's correlation coefficient, the following is an immediate corollary to Proposition 2.

Corollary 1. *The index of assortativity $a(p)$ of a two-pool assortative matching process with uniform assortativity F is constant with respect to p and equal to F.*

Statistical Identification of F

An empirical investigator will typically observe the proportions of the three possible types of matched pairs in a sample population. These can be used to estimate the parameters F and p under the maintained hypothesis that the matching was generated by a two-pool assortative process with uniform assortativity, F. The pairs that appear in a sample generated by a uniform two-pool assortative process will be draws from a multinomial distribution, where the probabilities of the three types of pairs are $(\pi_{11}, \pi_{12}, \pi_{22})$. The maximum likelihood estimators of these probabilities are equal to the sample proportions, $\hat{\pi}_{11}$, $\hat{\pi}_{12}$ and $\hat{\pi}_{22}$ of each of the three types of pairs. These estimates can then be used to identify estimates of the parameters p and F. The probability that a randomly drawn individual is of type 1 is equal to the probability that a randomly drawn pair of individuals has two type 1's plus one half the probability that a randomly drawn pair has one individual of each type, Therefore the probability p can be estimated as

$$\hat{p} = \hat{\pi}_{ii} + \frac{1}{2}\hat{\pi}_{12}v. \tag{5}$$

The probability π_{12} that a randomly drawn pair has one individual of each type must be $\pi_{12} = 2p(1-p)(1-F)$. Rearranging this expression to solve for F, and using our estimates $\hat{\pi}_{12}$ and $\hat{p}$, we construct an unbiased estimate $\hat{F}$ as

$$\hat{F} = 1 - \frac{\hat{\pi}_{12}}{2\hat{p}(1-\hat{p})} \tag{6}$$

The assumption that matching is a two-pool assortative process with uniform assortativity implies that $F \geq 0$. From Equations 5 and 6, we see that this requires that

$$\hat{\pi}_{12} \leq 2\left(\hat{\pi}_{11} + \frac{1}{2}\hat{\pi}_{12}\right)\left(1 - (\hat{\pi}_{11} + \frac{1}{2}\hat{\pi}_{12})\right) \tag{7}$$

Condition 7 simply requires that the observed fraction of "mixed" pairings be no larger than the expected fraction of mixed pairings if matching were entirely random.

Generalized Two-Pool Processes

The notion of two-pool assortative processes can be generalized to allow some types to be more likely to seek their match in an assortative pool than others. The process can also be generalized to allow there to be more than two types. Consider a population with n types, where p_i is the fraction of the population that is of type i. Suppose that with probability F_i an individual of type i obtains a partner from an *assortative pool* which includes only individuals of type i and with probability $G_i = 1 - F_i$, an individual obtains a partner by drawing randomly from a *random pool* which includes all individuals who have not joined assortative pools.[1]

[1] Of course this model is far from completely general. In this model, those who do not join an assortative pool consisting only of their own type, select their matches at random

The probability that a randomly selected matched pair includes one type i and one type j is equal to the sum of the probability that a randomly selected individual is a type i who has a type j partner and the probability that a randomly selected individual is a type j who has a type i partner. If $i \neq j$, a type i will have a type j partner only if the type i goes to the random pool rather than its assortative pool and happens to draw a type j. Since the random pool includes all individuals not in the assortative pools, the probability that a randomly drawn individual from the random pools is a type j must be

$$\frac{G_j p_j}{\sum_{k=1}^{n} G_k p_k}. \tag{8}$$

Therefore the probability that a randomly selected individual is a type i whose partner is a type j is

$$p_i G_i \left(\frac{G_j p_j}{\sum_{k=1}^{n} G_k p_k} \right) = \frac{p_i p_j G_i G_j}{\sum_{k=1}^{n} G_k p_k}. \tag{9}$$

This is also equal to the probability that a randomly selected individual is a type j who is matched with a type i. Therefore the probability that a randomly selected pair consists of one i and one j where $i \neq j$ is

$$\pi_{ij} = \frac{2 p_i p_j G_i G_j}{\sum_{k=1}^{n} p_k G_k}. \tag{10}$$

A type i will be matched with another type i either if it joins the assortative pool of its own type or if it joins the random pool and happens to draw another type i. The probability that a type i joins the assortative pool is $1 - G_i$. If it joins the random pool, the probability that it is matched with another type i is $p_i G_i / \left(\sum_{k=1}^{n} p_k G_k \right)$. It follows that

$$\pi_{ii} = p_i (1 - G_i) + \frac{p_i^2 G_i^2}{\sum_k p_k G_k}. \tag{11}$$

A randomly selected individual in the population could be drawn by first selecting a pair of individuals at random and then choosing one member at random from this pair. Thus it must be that or all i,

$$p_i = \pi_{ii} + \frac{1}{2} \sum_{j \neq i} \pi_{ij} \tag{12}$$

from a single pool that includes all individuals who did not join assortative pools. It does not allow the possibility, for example, that some members of types i and j, do not join assortative pools, but join a random pool that includes members of types i and j but no members of type k.

Non-Uniform Matching with Two Types: An Identification Problem

If there are only two types, then knowing the fraction of pairs that are of each possible composition is not sufficient to uniquely identify the two assortativity parameters, F_1 and F_2. To see this, let π_{11}, π_{12}, and π_{22} be the probabilities of each of the three possible compositions of mixed pairs. The probability that a randomly selected individual is of type 1 is then

$$p = \pi_{11} + \frac{1}{2}\pi_{12}. \tag{13}$$

Where $G_1 = 1 - F_1$ and $G_2 = 1 - F_2$, it follows from Equation 10 that

$$\frac{G_1 G_2}{pG_1 + (1-p)G_2} = \frac{\pi_{12}}{2p(1-p)}. \tag{14}$$

Equation 14 is satisfied when

$$G_1 = G_2 = \frac{\pi(1,2)}{2p(1-p)}. \tag{15}$$

From Equation 13, it follows that the probabilities π_{11}, π_{12}, and π_{22} uniquely determine the right side of Equation 15. Therefore the group composition π_{11}, π_{12} will be consistent with a two-pool matching process with uniform assortativity, $F = 1 - G > 0$ so long as $\pi_{12} < 2p(1-p)$, which is to say that the probability of a mixed match, is smaller than it would be if matching were random.

But the same data is consistent with a range of possible G_1, G_2 pairs where $G_1 \neq G_2$. Rearranging Equation, 14, we have

$$\frac{1}{pG_1} + \frac{1}{(1-p)G_2} = \frac{2}{\pi_{12}} \tag{16}$$

The set of assortativity parameters that are consistent with the distribution parameters π_{11}, π_{12}, π_{22} consists of all pairs of (G_1, G_2) that lie on the level curve defined by Equation 16 where $G_1 \leq 1$ and $G_2 \leq 1$. Simple calculations show that this is a downward-sloping curve with slope

$$\frac{dG_2}{dG_1} = -\left(\frac{1-p}{p}\right)\left(\frac{G_2^2}{G_1^2}\right) \tag{17}$$

Illustrative Diagrams

This section provides a brief diagrammatic explanation of the identification problem that arises when there are two types who have possibly differing degrees of assortativity. The downward-sloping curve in figure 1 shows the locus of combinations of assortativity parameters G_1 and G_2 for types 1 and 2 that are consistent with expected proportions $\pi_{11} = \pi_{22} = \pi_{12} = 1/3$ of paired combinations. In this case, the population fractions are $p = 1 - p = \pi_{11} + \frac{\pi_{12}}{2} = 1/2$. This curve intersects the upward-sloping 45 degree line at the point $G_1 = G_2 = G = 2/3$. Thus if there is uniform assortativity, the index of assortativity would be $F = 1 - G = 1/3$. In this case, each type would use the assortative pool with probability 1/3 and the random pool with probability 2/3. But the same proportions of types of matched pairs would be found with differing assortativities, F_1 and F_2, where $(F_1, F_2) = (1 - G_1, 1 - G_2)$ for any (G_1, G_2) pair that lies on the downward-sloping curve.

Figure 1: Assortativity Locus when $G_1 = G_2 = G = 2/3$.

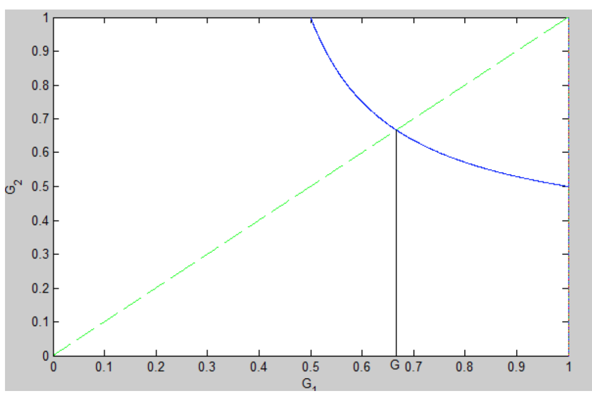
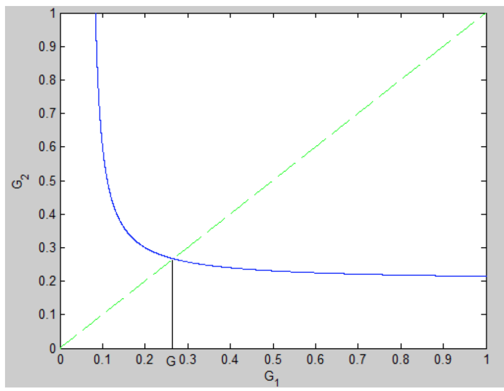


Figure 2 shows the locus of combinations of G_1 and G_2 that are consistent with the proportions $\pi_{11} = .7, \pi_{22} = .2, \pi_{12} = .1$. In this case the proportions of the two types in the population are $p = \pi_{11} + \pi_{12}/2 = .75$. In this case the curve intersects the upward-sloping 45 degree line at the point $G_1 = G_2 = G = .2667$. Thus if there is uniform assortativity, the index of assortativity would be $F = 1 - G = .7333$. Each type would use the assortative pool with probability .7333 and the random pool with probability .26678. The same proportions of matched pairs could also be achieved with differing assortativities by type as represented by any combination on the downward-sloping line. Although data on the composition of matching pairs is not sufficient to identify non-uniform assortativities when there are only two types, identification becomes possible when there are three or more types for which $0 < F_i < 1$.

Figure 2: Assortativity Locus when $\pi_{11} = .7, \pi_{22} = .2, \pi_{12} = .1$



Non-uniform Two-Pool Assortative Matching with Three Types

Proposition 3. *Consider a two-pool matching process with three types. Let F_i be the probability that a type i matches from an assortative pool consisting only of type i's and let $G_i = 1 - F_i$. Let π_{ij} be the probability that a randomly selected pair of individuals are of types i and j. There is a one-to-one mapping from the parameters $(G_1, G_2, G_3, p_1, p_2, p_3)$ of the matching process to the matching probabilities π_{ij}.*

The fact that the mapping from the assortativity parameters F_i to the probability distribution of compositions of matches is one-to-one means that we can estimate these parameters for a population by observing the actual distribution of match compositions in that population. The observed proportions will be maximum likelihood estimates of the probabilities of types of mixed matches. The estimates of these probabilities are in turn sufficient to provide maximum likelihood estimates of the assortativity parameters, conditional on the hypothesis that matching is a two-pool assortative process.

3 Strangers-in-the-Night Assortative Matching

In the two-pool assortative matching process, an individual who wishes to match assortatively can go to an "assortative pool" and be certain to obtain a match of the same type. We now consider an alternative matching process such that one is less likely to match assortatively, the less common is one's type. In particular, we assume that those who seek partners encounter others randomly with respect to their type. However, once they meet, the probability that two individuals will choose to form a partnership is higher if they are of the same type than if they are of different types. Because of the random nature of encounters, we call this the "strangers-in-the-night" process.

Strangers-in-the-Night with Uniform Assortativity

Definition 3. *A strangers-in-the-night matching process with two types and uniform assortativity is a matching process in which at all times a constant number of persons of each type are seeking partners. Those seeking partners are equally likely to meet a person of either type. When two persons of the same type meet, they match with probability s. When two persons of different types meet, they match with probability $m < s$.*

Let p and q be the fractions of those seeking partners who are of types 1 and 2 respectively. The conditional probabilities of one's partner's type, given one's own type are seen to be

$$\pi(1|1) = \frac{ps}{ps + qm} \tag{18}$$

and

$$\pi(1|2) = \frac{pm}{qs + pm}. \tag{19}$$

Therefore the degree of assortativity is

$$
\begin{aligned}
a(p) &= \left(\frac{ps}{ps+qm}\right) - \left(\frac{pm}{qs+pm}\right) \\
&= \frac{pq(s^2-m^2)}{(ps+qm)\,(qs+pm)} \\
&= \frac{pq(s^2-m^2)}{pq\,(s^2+m^2)+sm\,(p^2+q^2)} \\
&= \frac{pq(s^2-m^2)}{pq\,(s^2+m^2)+sm\,(1-2pq)} \\
&= \frac{pq(s^2-m^2)}{pq(s-m)^2+sm}. \qquad (20)
\end{aligned}
$$

Proposition 4. *If two types are matched by a strangers-in-the-night process with uniform assortativity, then Wright's F-statistic for the matching satisfies the equation*

$$
F(p) = \frac{pq(s^2-m^2)}{pq(s-m)^2+sm}, \qquad (21)
$$

If $s > m$, $F(p)$ is increasing in p for $p < 1/2$ and decreasing in p for $p > 1/2$, $F(0) = F(1) = 0$ and $F(1/2) = (s-m)(s+m)$.

Proof. According to Theorem 1, Wright's F-statistic is equal to the degree of assortativity. Therefore Equation 21 is immediate from Equation 20. Differentiating Equation 21, we find that if $s > m > 0$, then $a(p)$ is a strictly increasing function of $pq = p(1-p)$. This implies that $a(p)$ is maximized at $p = 1/2$ and is strictly increasing in p for $p < 1/2$ and strictly decreasing in p for $p > 1/2$. Direct calculations show that $a(0) = a(1) = 0$ and that $a(1/2) = (s-m)/(s+m)$. □

Statistical Identification of Parameters

In some applications, a researcher may be able to observe the makeup by type of existing matches, but would not have direct observations of the proportions of each type who seek matches. The proportions p and q of persons of each type seeking matches are not in general equal to the proportions found in existing matches. Since the more common type is more likely to meet its own type than is the less common type, the rate at which the more common type finds matches will exceed that for the less common type. The rates at which type 1's and type 2's form partnerships are respectively $ps + qm$ and $pm + qs$, and the difference between these two rates is $(p - q)(s - m)$. [2] However, if we assume that matching is the result of a strangers-in-the-night process with uniform assortativity, it is possible to estimate the proportions p and q as well as the ration m/s from observations of the makeup of matches that form.

The rate at which two type 1's meet and form partnerships is $p^2 s$. The rate at which two type 2's meet and form partnerships is $q^2 s$ and the rate at which mixed partnerships of one type 1 and one type 2 meet and form partnerships is $2pqm$. Thus the total rate at which partnerships are formed is $(p^2+q^2)s+2pqm$. The expected ratios of partnership types in the population will therefore be

$$\pi_{11} = \frac{p^2 s}{(p^2 + q^2)s + 2pqm} \tag{22}$$

$$\pi_{22} = \frac{q^2 s}{(p^2 + q^2)s + 2pqm} \tag{23}$$

$$\pi_{12} = \frac{2pqm}{(p^2 + q^2)s + 2pqm} \tag{24}$$

Although we do not directly observe p and $q = 1 - p$, we can estimate these parameters from the observed fractions of pairs of each type. From Equations 22 and 23, it follows that

$$\frac{p}{q} = \sqrt{\frac{\pi_{11}}{\pi_{22}}}. \tag{25}$$

Therefore it follows that if $\hat{\pi}_{ij}$ is the observed fraction of pairs that are of type ij we can estimate the fraction p by

$$\hat{p} = \frac{\sqrt{\hat{\pi}_{12}}}{\sqrt{\hat{\pi}_{11} + \hat{\pi}_{22}}} \tag{26}$$

From Equations 22, 24, and 25 it follows that

$$\frac{m}{s} = \frac{p}{2q}\frac{\pi_{12}}{\pi_{11}} = \frac{\pi_{12}}{2\sqrt{\pi_{11}\pi_{22}}} \tag{27}$$

[2] To make a stationary model of this process, we need individuals to have lives of finite length. Some individuals of each type reach the end of their life without finding a match. Given its lower matching probability, the less common type will be more likely than those of the more common type to die without finding a match.

Therefore we can estimate the ratio m/s by

$$\widehat{\left(\frac{m}{s}\right)} = \frac{\hat{\pi}_{12}}{2\sqrt{\hat{\pi}_{11}\hat{\pi}_{22}}}. \tag{28}$$

Although the proportions $\hat{\pi}_{ij}$ are sufficient to identify estimates of p, q, and m/s, they are not sufficient to allow separate estimates of m and s. To see this, we note that if p, m, and s satisfy Equations 22-24, then p, km and ks would also satisfy these equations for any $k > 0$.

From Equation 20 it follows that if $s > m > 0$, then $a(p)$ is a strictly increasing function of $pq = p(1-p)$. Therefore $a(p)$ is maximized at $p = 1/2$ and is strictly decreasing in p for $p > 1/2$ and strictly decreasing in p for $p < 1/2$. Simple calculations show that $a(1/2) = (s-m)/(s+m)$ and $a(0) = a(1) = 0$.

Strangers-in-the-night with Non-uniform Assortativity

It might be that the two types who meet by a strangers-in-the-night process differ in the probabilities of matching with another individual of the same type. Suppose that the probability that an encounter between two type i's leads to a partnership is s_i for $i = 1, 2$. Then it must be the case that

$$\frac{\pi_{11}}{\pi_{22}} = \frac{s_1}{s_2}\left(\frac{p}{q}\right)^2. \tag{29}$$

In order to identify the ratio s_1/s_2, we would need to have independent estimates of the proportions p and q of those seeking matches who are of each type. Identification of separate values for s_1 and s_2 would be possible if and only if direct information of the proportions p and q were available.

4 Dynamics and Assortativity

William Hamilton's [6] theory of altruistic behavior among kin relatives is one of the most celebrated applications of Wright's F-statistic. Hamilton presented this theory by considering simple symmetric interactions between pairs of related animals who could help each other at a cost to themselves. Hamilton's helping game as a game in which there are two players, each of whom can increase the expected reproductive success (fitness) of the other at some cost to itself.

Definition 4. *Hamilton's helping game is a two player game in which a strategy for either player is a level of effort $x \in [0,1]$ that this player can* *exert to help the other. Where x_i is the strategy of player i, the reproductive fitness of player 1 is $b(x_2) - c(x_1)$ and that of player 2 is $b(x_1) - c(x_2)$. There are positive but diminishing marginal benefits from help received and increasing marginal costs to help given. Thus we assume that $b(0) = c(0) = 0$, $b'(x) > 0$, $c'(x) > 0$, $b''(x) < 0$ and $c''(x) > 0$ for all $x \in [0, 1]$.*

Evolutionary Equilibrium with two-pool assortative matching

Suppose that animals are matched by a two-pool assortative matching process with assortativity F to play Hamilton's helping game. Types are distinguished by the amount x of help that they will offer to their partners. With probability F, an animal of type x is matched with an animal of its own type from the assortative pool and with probability $1-F$ it is matched with a partner drawn randomly from the population distribution.

The expected fitness of type x is

$$Fb(x) + (1 - F)y - c(x) \tag{30}$$

where y is the expected value of benefits received if one is matched with a random selection from the population distribution. An animal of another type x' will receive $b(x')$ if it joins the assortative pool of its own type and will face the same probability distribution of benefits as the x type if it matches from the random pool. The expected fitness of a type x will exceed that of a type x' if

$$b(x)F - c(x) > b(x')F - c(x'). \tag{31}$$

Given our assumptions on b and c, there will be a unique $\bar{x}$ such that $\bar{x}$ maximizes the function $b(x)F - c(x)$ on the interval $[0, 1]$. Then regardless of the proportions of types in the population, animals of type $\bar{x}$ will have greater fitness than any other type.

Proposition 5. *If animals are matched by a two-pool assortative process with assortativity F, where type is determined by the amount, x, that each will offer in a Hamilton helping game with its partner, and if the reproduction rate of each type is an increasing function of its fitness, then there is a unique evolutionary equilibrium in which all animals are of type $\bar{x}$ where $\bar{x}$ maximizes $b(x)F - c(x)$.*

Proposition 5 predicts that in equilibrium, all players will choose the strategy of offering $\bar{x}$ units of health where the ratio of the marginal cost of helping to the marginal benefit received is equal to F. This condition is consistent with Hamilton's description of equilibrium for interaction of animals that are genetically related:

> "The social behavior of a species evolves in such a way that in each distinct behavior-evoking situation the individual will seem to value his neighbors fitness against his own according to the coefficients of relationship appropriate to that situation." [6]

Evolutionary equilibrium with strangers-in-the-dark

If matching to play Hamilton's helping game takes place assortatively according to a strangers-in-the-dark process rather than a two-pool process, then instead of a unique stable equilibrium, there will be a continuum of stable equilibria. These equilibria will be characterized by having a positive fraction of the population who devote the same positive effort to helping and a complementary fraction of the population who offer no help. More outcomes are stable because, with this process, rare mutants that do well when they encounter their own type are highly unlikely to meet their own type. In contrast, if matching occurs by means of a two-pool process with assortativity F, a rare mutant has a probability of at least F of being matched with one of its own type.

Definition 5. *Define the cost-benefit ratio function for a Hamilton's helping game as* $\rho(x) = c(x)/b(x)$. *A Hamilton helping game is* regular *if* $\rho(\cdot)$ *is a continuous, increasing function on the interval* $(0, 1]$ *with* $\rho_0 = \lim_{x\to 0} \rho(x) \geq 0$.

The following proposition, which is proved in the appendix to this paper, asserts that there is a continuum of evolutionarily stable equilibria for a regular Hamilton's helping game in which partners are matched by a strangers-in-the-night process.

Proposition 6. *Suppose that individuals are matched by a strangers-in-the-dark process to play a regular Hamilton's helping game, where* $\rho_0 < \frac{s-m}{s+m} < \rho(1)$. *Let* $x^* = \rho^{-1}\left(\frac{s-m}{s+m}\right)$. *Then for every* x *in the interval* $(0, x^*]$, *there is a unique* $p(x) \in [1/2, 1)$ *such that there is a locally stable equilibrium in which the fraction* $p(x)$ *of those seeking partners offer* x *units of help and the fraction* $1 - p(x)$ *offer no help. The fraction* $p(\cdot)$ *is a decreasing function of* x.

5 Appendix

Proof of Proposition 1

Proof. To motivate Wright's F-statistic as a correlation, let us construct two random variables I_A and I_B as follows. If one randomly selects one matched pair and then randomly chooses one individual from that pair, let I_A be the random variable that takes on the value 1 or 0 depending on whether this individual is a type 1 or a type 2. Let I_B be the random variable that takes the value 1 or 0 depending on whether the remaining member of the selected pair is of type 1 or type 2. Wright's F is the correlation coefficient between the random variable I_A and I_B. This correlation coefficient is, by definition,

$$\rho = \frac{E(I_A I_B) - E(I_A)E(I_B)}{\sigma(I_A)\sigma(I_B)} \tag{32}$$

Now $E(I_A) = E(I_B) = p$, and $\sigma(I_A) = \sigma(I_B) = \sqrt{p(1-p)}$. Also $E(I_A I_B) = \pi_{11} = p\pi(1|1)$. Therefore Equation 32 can be written as

$$\rho = \frac{p\pi(1|1) - p^2}{p(1-p)} = \frac{\pi(1|1) - p}{1-p} \tag{33}$$

This establishes Equation 1.

Since $\pi(1|1) = 1 - \pi(2|1)$, Equation 33 can be written as

$$\rho = \frac{1 - \pi(2|1) - p}{1-p} = 1 - \frac{\pi(2|1)}{1-p} \tag{34}$$

Then, since $\pi(1,2) = p\pi(2|1)$, it follows that

$$\rho = 1 - \frac{\pi_{12}}{p(1-p)} \tag{35}$$

This establishes Equation 2.

Since $\pi_{12} = p\pi(2|1) = (1-p)\pi(1|2)$, it must be that

$$\pi(1|2) = \frac{p}{1-p}\pi(2|1) = \frac{p}{1-p}\left(1 - \pi(1|1)\right). \tag{36}$$

From Equation 1 it follows that $\pi(1|1) = (1-p)F(p)+p$. Therefore Equation 36 simplifies to

$$\pi(1|2) = p\left(1 - F(p)\right). \tag{37}$$

Therefore we have

$$\begin{aligned} \pi(1|1) - \pi(1|2) &= (1-p)F(p) + p - p\left(1 - F(p)\right) \\ &= F(p) \end{aligned} \tag{38}$$

This establishes Equation 3.

$\square$

Proof of Proposition 3

Proof. The mapping from the G_i's and p_i's to the probabilities π_{ij} is immediate from Equations 10 and 11.

To find the inverse mapping from the π_{ij}'s to the G_i's and p_i's, we proceed as follows. From Equation 10, it follows that

$$\frac{2\,(p_1p_2G_1G_2)\,(p_1p_3G_1G_3)}{p_2p_3G_2G_3} = \frac{\pi_{12}\pi_{13}}{\pi_{23}}\left(\sum_{k=1}^{3} p_kG_k\right) \tag{39}$$

Simplifying and rearranging Equation 39, we have

$$p_1G_1 = \frac{1}{\sqrt{2}}\frac{\sqrt{\pi_{12}\pi_{13}}}{\sqrt{\pi_{23}}}\sqrt{\sum_k p_kG_k} = \frac{1}{\sqrt{2}}\frac{\sqrt{\pi_{12}\pi_{23}\pi_{13}}}{\pi_{23}}\sqrt{\sum_k p_kG_k}. \tag{40}$$

Symmetric reasoning shows that also

$$p_2G_2 = \frac{1}{\sqrt{2}}\frac{\sqrt{\pi_{12}\pi_{23}\pi_{13}}}{\pi_{13}}\sqrt{\sum_k p_kG_k} \tag{41}$$

$$p_3G_3 = \frac{1}{\sqrt{2}}\frac{\sqrt{\pi_{12}\pi_{23}\pi_{13}}}{\pi_{12}}\sqrt{\sum_k p_kG_k} \tag{42}$$

Summing the terms in Equations 40-42, we have

$$\sum_k p_kG_k = \frac{1}{\sqrt{2}}\sqrt{\pi_{12}\pi_{23}\pi_{13}}\sqrt{\sum_k p_kG_k}\left(\frac{1}{\sqrt{\pi_{23}}+\sqrt{\pi_{13}}+\sqrt{\pi_{12}}}\right) \tag{43}$$

which in turn implies

$$\sqrt{\sum_k p_kG_k} = \frac{1}{\sqrt{2}}\sqrt{\pi_{12}\pi_{23}\pi_{13}}\left(\frac{1}{\sqrt{\pi_{23}}+\sqrt{\pi_{13}}+\sqrt{\pi_{12}}}\right) \tag{44}$$

From Equations 40-42 it then follows that

$$p_1G_1 = \frac{1}{2}\pi_{12}\pi_{13}\left(\frac{1}{\sqrt{\pi_{23}}+\sqrt{\pi_{13}}+\sqrt{\pi_{12}}}\right) \tag{45}$$

$$p_2G_2 = \frac{1}{2}\pi_{12}\pi_{23}\left(\frac{1}{\sqrt{\pi_{23}}+\sqrt{\pi_{13}}+\sqrt{\pi_{12}}}\right) \tag{46}$$

$$p_3G_3 = \frac{1}{2}\pi_{13}\pi_{23}\left(\frac{1}{\sqrt{\pi_{23}}+\sqrt{\pi_{13}}+\sqrt{\pi_{12}}}\right) \tag{47}$$

Since

$$p_i = \pi_{ii} + \frac{1}{2}\sum_{j \neq i} \pi_{ij} \tag{48}$$

the p_i's are uniquely determined by the π_{ij}'s. Given that the p_i's are uniquely determined, it follows from Equations 40-42 that the G_i's are also uniquely determined by the π_{ij}'s. This proves Proposition 3.

□

Proof of Proposition 6

Proof. The function $F(\cdot)$ is continuous and strictly decreasing for p in the interval $[1/2, 1]$, with $F(1/2) = \frac{s-m}{s+m}$ and $F(1) = 0$. Therefore $F^{-1}(\cdot)$ is a continuous, decreasing function from the interval $[0, \frac{s-m}{s+m}]$ onto $[0, 1/2]$. Our assumptions imply that function $\rho(\cdot)$ is a continuous, increasing function from $(0, x^*]$ onto the interval $(\rho_0, \frac{s-m}{s+m}]$. Therefore there is a well-defined function $p(x) = F^{-1}(\rho(x))$ mapping the non-empty interval $(0, x^*]$ onto $[1/2, 1)$. Since $\rho(\cdot)$ is an increasing function and $F^{-1}(\cdot)$ is a decreasing function, it must be that $p(\cdot)$ is a decreasing function of x.

Let $x \in (0, x^*]$ and suppose that the fraction $p(x)$ of the population seeking matches are of the type that contributes x and the fraction $1-p(x)$ are of the type that contributes 0. Then the expected payoff of a type x is $\pi(x|x)b(x) - c(x)$ and the expected payoff of a type 0 is $\pi(x|0)b(x)$. The difference between the expected payoffs of the two types is $F(p(x)b((x) - c(x)$. From the definition of $p(x)$, it follows that $F(p(x)) = \rho(x)$ and hence $b(x)F(p(x)) - c(x) = 0$. Therefore, when the fraction $p(x)$ are of type x and fraction $1-p(x)$ are of type 0, the expected fitnesses of the two types are equal. A mutant individual of another type, who contributes a non-zero amount $x' \neq x$ will almost always meet either a type x or a type 0. The probability that a type x' matches with a type x is the same as the probability that a type 0 matches with a type x. Therefore the fitness of a type x' is $\pi(x|0)b(x) - c(x') < \pi(x|0$, which is the expected payoff of the two incumbent types x and 0.

□

References

[1] Ingela Alger. Public goods games, altruism, and evolution. *Journal of Public Economic Theory*, 12(4):789–813, 2008.

[2] Ingela Alger and Jörgen W. Weibull. Kinship, incentives and evolution. *American Economic Review*, 100(4):1725–1758, 2010.

[3] Ingela Alger and Jörgen W. Weibull. Homo moralis: Preference evolution under incomplete information and assortative matching. `http://ideas.repec.org/p/tse/wpaper/25607.html`, 2012.

[4] Theodore Bergstrom. The algebra of assortative mating and the evolution of cooperation. *International Game Theory Review*, 5(3):1–18, 2003.

[5] L.L. Cavalli-Sforza and M.W. Feldman. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton, N.J., 1981.

[6] W.D. Hamilton. The genetical evolution of social behavior, parts i and ii. *Journal of Theoretical Biology*, 7:1–52, 1964.

[7] Daniel H. Hartl and Andrew G. Clark. *Principles of Population Genetics*. Sinauer Associates, Sunderland, Massachusetts, 1989.

[8] Sewall Wright. Systems of mating. *Genetics*, 6:111–178, 1921.

[9] Sewall Wright. Coefficients of inbreeding and relationship. *The American Naturalist*, 56(645):330–338, Jul-Aug 1922.

[10] Sewall Wright. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, 19(3):395–420, September 1965.