

7 K

The Evolution of Families

From an evolutionary perspective, the family is a hostage situation. Offspring are gametes (eggs and sperm) who run away together with stolen copies of parental genes. Their stolen goods put offspring in a position of power. A genetic variant in a parent that hurts offspring very likely hurts a copy of itself. A genetic variant that helps offspring very likely helps a copy of itself. This is not an entirely cooperative scenario, because offspring are not their parents. They do not have exactly the same interests. They may use their power to extract more help than they really need. “Give me another worm,” the baby bird cries, “or your lineage gets it!”

Animal societies scaffold on this hostage scenario. Everywhere in nature, common descent is an important source of **positive assortment**. But this engenders always a mix of conflict and cooperation. If there is cooperation within the family, there is also competition between families. In many mammals for example, social groups are sisterhoods. And these sisterhoods compete intensely with one another for space, food, and a safe future for their offspring. Sisters also compete with one another. And the balance between conflict and cooperation can tip suddenly to either side. The baboon says: *I against my sister. I and my sister against my cousin. I, my sister, and my cousin against the world.*¹

Human societies build absurd and amazing institutions on top of hostages. In medieval Europe, royalty dominated political power. And royalty is a family. But royal families routinely raised one another’s children, literal royal hostages, to ensure that they could remain at peace. “Don’t invade my lands,” the princess cries, “or your princess gets it!”²

This chapter is about **kin selection**.

Hamilton’s rule

When relatives interact, it can promote cooperation. Under strong assumptions, there is a simple rule.

¹This is a slightly modified proverb (Nafisa Haji, The Sweetness of Tears).

²Peacemaking in Middle Ages citation

Suppose two relatives interact in a one-shot **prisoner's dilemma**. When will selection favor cooperation instead of non-cooperation? Countless students have been taught in introductory evolutionary biology and anthropology courses that there is a simple answer to this question:

$$rb > c$$

- 5 where r is the **coefficient of relatedness** between the pair, and b and c have their usual meanings in the prisoner's dilemma. This condition is usually called **Hamilton's rule**.³ This rule is a valuable qualitative result. It tells us that recent common descent, which increases r , can increase the range of conditions under which natural selection favors cooperation. The model it is derived from also helps us
10 understand why.

Here is a very simple and sloppy derivation. The rest of this section derives the rule in a more careful and educational way. Recall the statistical positive assortment model from **5C**. In this model, the probability an individual of any strategy X is paired with another individual with the same strategy is:

$$\Pr(X|X) = r + (1 - r)p$$

- 15 where r is a coefficient of assortment and p the proportion of the population with strategy X. Now imagine pairs of individuals interact once in the additive prisoner's dilemma from the previous chapter. The condition for rare cooperative (C) individuals to invade a population of non-cooperators (N) is:

$$\Pr(C|C)(b - c) + \Pr(N|C)(-c) > 0$$

- The left side of the condition above is the expected fitness of the rare C individual.
20 The right side is the average fitness of a common N. Now we fill in the abstract interaction probabilities with the statistical assortment model:

$$(r + (1 - r)p)(b - c) + (1 - r)(1 - p)(-c) > 0$$

When C is rare, $p \approx 0$, so this simplifies to $rb - c > 0$. This is Hamilton's rule.

- Like most results in evolutionary game theory, this rule is a result of very specific set of assumptions. So it is not completely general. But there is a tendency
25 to exaggerate its generality, especially in the human evolutionary sciences. And this leads to mistakes. For example, the founders of Evolutionary Psychology wrote that it "should be true of any species on any planet at any time."⁴ But consider the following two puzzles.

- 30 In many species of bird, offspring do not sexually mature in the first year but can instead stay behind and help their parents at the nest. Suppose a situation arises in which the helper can sacrifice her own life to save four siblings. So $b = 4$,

³cite Hamilton

⁴Page 168 of Berkow, Cosmides, Tooby (1992) The Adapted Mind. It is tempting to dismiss this, because few people expect evolutionary psychologists to have a strong grasp of evolutionary theory. But many biologists are equally confused. They are just less likely to put it in writing.

$c = 1$, and $r = 0.5$. Hamilton's rule is satisfied, but any individual in which a gene arises that leads to this self-sacrifice would die and remove the allele from the population. Why does Hamilton's rule arrive at the wrong answer?

The coefficient of relatedness r is $1/2$ for full siblings, $1/4$ for cousins, and so on, declining with genealogical distance. This coefficient is often interpreted as the probability a pair of individuals share an allele. But this can't be right, because at many sites in the genome, in most species, there are only a few alleles. So many distantly related individuals will share alleles. Instead r is the probability the individuals share an allele by common descent. They must have inherited it from a common ancestor. But again, this can't be right. If mutations are rare, and a population is not very large, then most individuals who share an allele must share it by common descent. In such populations, why doesn't natural selection favor universal cooperation? Or why doesn't Hamilton's rule predict that it should, if indeed r is the probability of sharing an allele by common descent?

The Price equation. To sort out Hamilton's rule, we need to back up and start over. We need a more general statement of natural selection. Then we can introduce assumptions, until we get the rule back. The place to start is the **Price equation**:

$$\bar{w}\Delta p = \text{cov}(w_i, p_i)$$

This equation states a simple fact about evolution: The change in the proportion of the population p that has any particular allele is equal to the covariance between individual fitness w_i and individual allele frequency p_i . I show you how to derive this equation in the box on the next page.

The Price equation is a very general statement about evolutionary change. It isn't about natural selection, specifically. If w_i is random, then it's a neutral model of evolutionary change. It is also perfectly general to a wide variety of inheritance systems, which is important when modeling some social insects for example, where females are diploid and males haploid.

Now what we'll do is make one assumption at a time until we recover Hamilton's rule.⁵

Additive fitness. To start, we assume an additive prisoner's dilemma as the cooperation game. Remember, there is nothing necessarily additive about cooperation payoffs. Assume:

$$w_i = \omega + h_i(-c) + y_i b$$

The symbol ω is baseline fitness unrelated to the cooperation game. The variables b and c have their usual meanings. Now h_i is the probability that individual i provides help, and y_i is the probability that same individual receives help from its partner.

⁵This derivation is a simplified version of Grafen's geometric view of relatedness article.

Deriving the Price equation. We focus on a particular allele “S” and derive an expression for its expected change in frequency in the population. Let n_i be the number of copies of any allele that individual i could possess, its *ploidy*. So for example, humans are $n = 2$ for most chromosomes. Let N be the size of the allele population, $N = \sum_i n_i$. Let p_i be the proportion of individual i ’s alleles that are S. Then just by definition, the proportion of S in the population is:

$$p = \sum_i \frac{n_i p_i}{N}$$

Now we use a prime symbol $'$ to indicate the value of a variable in the next generation. So the proportion of S in the next generation is just $p' = \sum_i (n'_i p'_i) / N'$. Now let w_i be the “fitness” of individual i , the number of offspring it produces. Let $\bar{w} = \sum_i (n_i / N) w_i$ be the average fitness. Then we can substitute $n'_i = w_i n_i$ and $N' = \bar{w} N$.

$$p' = \sum_i \frac{w_i n_i p'_i}{\bar{w} N}$$

So far, nothing but definitions of random variables. Let’s make the above into a difference equation, instead of a recursion, $\Delta p = p' - p$:

$$\Delta p = \sum_i \frac{w_i n_i p'_i}{\bar{w} N} - p$$

Multiplying both sides by \bar{w} :

$$\bar{w} \Delta p = \sum_i \frac{n_i}{N} w_i p'_i - p \bar{w} = \sum_i \frac{n_i}{N} w_i p'_i - p \sum_i \frac{n_i}{N} w_i$$

Now just for fun, let’s add and subtract the term $\sum_i (n_i / N) w_i p_i$ to the right side. This doesn’t change the value, because we added it and subtracted it at the same time. But then we can arrange and factor to make a simpler expression. Like this:

$$\bar{w} \Delta p = \underbrace{\sum_i \frac{n_i}{N} w_i p'_i}_A - \underbrace{p \sum_i \frac{n_i}{N} w_i}_B + \underbrace{\sum_i \frac{n_i}{N} w_i p_i}_C - \underbrace{\sum_i \frac{n_i}{N} w_i p_i}_D$$

The terms A and B are the originals. We added C and subtracted D. Now let’s rearrange:

$$\begin{aligned} \bar{w} \Delta p &= \left(\underbrace{\sum_i \frac{n_i}{N} w_i p_i}_C - \underbrace{\sum_i \frac{n_i}{N} w_i}_B \right) + \left(\underbrace{\sum_i \frac{n_i}{N} w_i p'_i}_A - \underbrace{\sum_i \frac{n_i}{N} w_i p_i}_D \right) \\ &= \sum_i \frac{n_i}{N} w_i (p_i - p) + \sum_i \frac{n_i}{N} w_i (p'_i - p_i) \end{aligned}$$

Now the last step. The definition of covariance is $\text{cov}(x, y) = E(x(y - \bar{y}))$. So the term on the left above is the covariance between w_i and p_i , since p is the average p_i . And the term on the right above is an average. Using these definitions, we arrive at the Price equation:

$$\bar{w} \Delta p = \text{cov}(w_i, p_i) + E(w_i \Delta p_i)$$

The term $\Delta p_i = p'_i - p_i$ is the change within individual i . In most cases, this is zero, because the frequency of an allele does not change within an individual within its lifetime. But it’s not impossible. Setting $\Delta p_i = 0$ gives us the simplified form in the text.

Additive fitness is a very special case that disallows synergies and diminishing returns. For most resources, if you give me twice as much, my benefit is not twice as large.

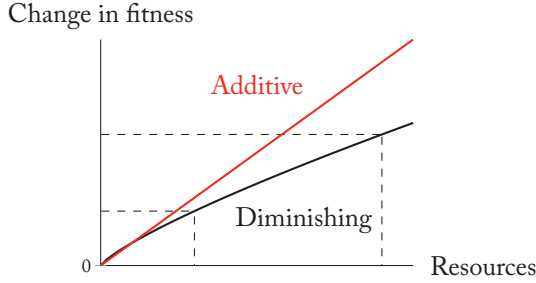


Figure
7.1

Now insert our definition for w_i into the Price equation:

$$\bar{w}\Delta p = \text{cov}(\omega + h_i(-c) + y_i b, p_i)$$

Now since a covariance is an expectation, and an expectation is a sum, we can decompose the above across addition, giving us a sum of three covariances:

$$\bar{w}\Delta p = \text{cov}(\omega, p_i) + \text{cov}(h_i(-c), p_i) + \text{cov}(y_i b, p_i)$$

The first covariance, $\text{cov}(\omega, p_i)$, is zero. The reason is that ω is a constant, and constants don't vary, so they also don't covary. This leaves the second and third covariances. For these, we factor out the constants:

$$\bar{w}\Delta p = (-c) \text{cov}(h_i, p_i) + b \text{cov}(y_i, p_i)$$

And this is greater than zero, meaning cooperation increases, when:

$$b \frac{\text{cov}(y_i, p_i)}{\text{cov}(h_i, p_i)} > c$$

This is not yet Hamilton's rule. Not even close. But it does resemble it. Instead of r , we have the ratio of the covariance between receiving help and genotype to the covariance between giving help and genotype.

Additive genetics. Next we need to specify how genotype p_i relates to phenotypes h_i and y_i . We make each phenotype an additive function of allele frequency p_i .

$$h_i = a + kp_i$$

$$y_i = a + kp_j$$

where p_j is the allele frequency in the partner j . There are other possibilities. Many alleles are dominant, and so two copies and one copy produce the same phenotype. This model doesn't allow that. Also no recessive alleles. Each additional cooperation allele produces a dose k of helping.

We insert these definitions into our rule and simplify:

$$b \frac{\text{cov}(a + kp_j, p_i)}{\text{cov}(a + kp_i, p_i)} > c$$

As before, constants don't vary and so they don't covary, so the a terms vanish. This leaves us:

$$b \frac{k \text{cov}(p_j, p_i)}{k \text{cov}(p_i, p_i)} > c$$

The covariance of a variable with itself is its variance, and so we arrive at:

$$b \frac{\text{cov}(p_j, p_i)}{\text{var}(p_i)} > c$$

5 This is usually expressed as:

$$b\beta(p_j, p_i) > c$$

where $\beta(p_j, p_i)$ is the regression coefficient for the line regressing the partner's genotype p_j on the focal individual's p_i . In other words, it expresses how the focal's genotype predicts the partner's genotype.

This implies there is linear equation that satisfies:

$$\begin{aligned} E(p_j|p_i) &= E(p_j) + \beta(p_j, p_i)(p_i - E(p_i)) \\ &= p + \beta(p_j, p_i)(p_i - p) \end{aligned}$$

10 Solving for $\beta(p_j, p_i)$:

$$\beta(p_j, p_i) = \frac{E(p_j|p_i) - p}{p_i - p}$$

Okay, this still isn't r . But it is a function of the genotypes of the two individuals. Any process that makes $E(p_j|p_i) > p$ would support the evolution of cooperation in this model. But what does relatedness have to do with it?

15 **Very weak selection.** We need a way to make $\beta(p_j, p_i) = r$. And it turns out that as we let the strength of natural selection become weak, $\beta(p_j, p_i) \approx r$. We'll proceed by constructing an expression for $E(p_j|p_i)$, the expected genotype of j given i , and relating it to shared ancestry. Then we'll see why our expression requires weak selection to be valid.

20 Define r as the proportion of genes that are **identical by descent** (IBD) in individuals i and j . When a gene is IBD, it means i and j have the same allele and they got it from a common ancestor. Now when an allele at the cooperation locus is IBD, it may or may not also be an allele that increases cooperation. Suppose for example individual i has $p_i = 0.5$, which means perhaps a diploid species and one copy of the cooperation allele. So while i and j share overall r alleles by common
25 descent, there's only a half chance it's a cooperation allele, because i only has half

cooperation alleles. This means the chance j shared a cooperation allele with i through common descent is rp_i .

The pair may also share cooperation alleles through the non-IBD part of their genomes, which is of size $1 - r$. For that part of the genome, the expected proportion of cooperation alleles is the population average, p . So the chance j shares a cooperation allele with i that is not IBD is $(1 - r)p$.

Together we get:

$$E(p_j|p_i) = \underbrace{rp_i}_{\text{IBD}} + \underbrace{(1-r)p}_{\text{not-IBD}}$$

Now solve for r :

$$r = \frac{E(p_j|p_i) - p}{p_i - p}$$

And this is equal to $\beta(p_j, p_i)$, the regression from the previous section.

Okay, so where's the weak selection? It's already in there, smuggled in under shadows. When I asserted that the proportion of cooperation alleles that are IBD is rp_i , I implicitly assumed that cooperation and non-cooperation alleles have the same chances of being passed down from a common ancestor. But if there is selection, this can't be true. The fact that you exist implies for example that you didn't inherit any lethal alleles from any ancestor. So if selection has favored cooperation (or non-cooperation) in past generations, then inheritance won't be fair in the way we usually calculate r , by just counting links in the pedigree. Technically we need to assume very weak selection (neutrality) to do that.

For close relatives, selection doesn't have to be neutral for r to be a good approximation. But for distant relatives, even weak selection will make r a terrible device for predicting shared alleles. Once an allele goes to fixation in a population, then almost everyone shares it by common descent. But that isn't what's relevant to understanding the evolution of cooperation. What's relevant is knowing what happens to a mutant allele. Close relatives tend to share recent mutations, but only if those mutations are lethal or too strongly selected against. And that creates positive assortment of phenotypes among close relatives.

Okay, if we are comfortable letting selection be weak, then we have now recovered Hamilton's rule. If you make other assumptions about the game or about genetics or about the strength of selection, you get a different rule. At least on this planet.

Making kin selection models

The lack of generality in Hamilton's rule means you shouldn't use it to make models. It is a result of a very specific combination of assumptions. What you should do instead is specify how social groups form, whether they are comprised of relatives, and then derive the implications of those assumptions.

This has to be emphasized, because sometimes biologists attempt an analysis by writing an **inclusive fitness** expression for each strategy. Inclusive fitness in this context means an individual's own personal payoffs plus a fraction r of a relative's payoffs. This by analogy to Hamilton's rule, because it's as if the individual receive
5 $-c + rb$, with the benefits to the partner discounted by relatedness r .

Consider as an example the Hawk-Dove game from **4H**. The Hawk-Dove game is not additive in its payoffs. To remind you:

	Hawk	Dove
Hawk	$(v - c)/2$	v
Dove	0	$v/2$

What happens depends upon what both players do. So instead of trying to write
10 some additive inclusive fitness expression, and possibly botching it, just use the statistical assortment model. The fitness of Hawk is:

Reproductive skew

x

What is inclusive fitness and does selection maximize it?

15 x

Group selection

x [or do this in its own chapter?]

Neighbor-modulated fitness

x [put this in societies chapter maybe?]