The Price Equation
2015-03-24 2:50pm

Abstract

The Price equation is a mathematical description of evolution. Its abstract nature allows it to be applied to evolutionary processes of different kinds, both biological and nonbiological.

My aim here is to give, in one place, simple and rigorous derivations of Price’s equation and the criteria for kin and group selection. The equivalence of kin and group selection follows immediately.

\[ \text{Evolution} = \text{Selection} + \text{Transmission Bias} \]

Assign a number \( z \) to each member of a set \( S \). Let \( \bar{z} \) be the \( z \)-average over members of \( S \).

Partition \( S \) into mutually exclusive and exhaustive subsets \( S_i \). Let \( Z_i \) be the \( z \)-average in \( S_i \). Set \( \bar{Z} = \sum q_i Z_i \), a weighted average of the \( Z_i \). We still use \( \bar{z} \) when it is regarded as an average of \( z \)-values of members of \( S \).

Partition a second set \( S' \) into an equal number of subsets \( S'_i \), with analogous quantities \( z', \bar{z}', Z'_i, q'_i \). In general \( |S'| \neq |S| \).

Set \( \Delta \bar{z} = \bar{z'} - \bar{z}, \Delta Z_i = Z'_i - Z_i, \) and \( \Delta q_i = q'_i - q_i \).

**Theorem** (Price equation).

\[ \Delta \bar{z} = \sum \Delta q_i Z_i + \sum q'_i \Delta Z_i, \]  

(1)

The terms in the equation have intuitive names:

- \( \Delta \bar{z} \) is the **evolution** of \( \bar{z} \) from \( S \) to \( S' \).
- \( \sum \Delta q_i Z_i \) is the **selection** part of the evolution of \( \bar{z} \). It measures the “success” of \( z \) from \( S \) to \( S' \) in terms of the averages \( Z_i \).
- \( \sum q'_i \Delta Z_i \) is the **transmission bias** part of the evolution of \( \bar{z} \). It is a weighted average of the \( z \)-average changes from the \( S_i \) to the \( S'_i \).

**Proof:**

\[ \Delta \bar{z} = \sum q'_i Z'_i - \sum q_i Z_i = \sum (q'_i - q_i)Z_i + \sum q'_i(Z'_i - Z_i) \]
\[ = \sum \Delta q_i Z_i + \sum q'_i \Delta Z_i. \]

Price’s equation is a trivial algebraic identity (a tautology). But it usefully separates the evolution \( \Delta \bar{z} \) into two parts in a very general setting.

---

1 Notation: \( |S| \) is the number of members of the set \( S \).
**Covariance Form**

Set $W_i = |S'_i|/|S_i|$, the fitness of $S_i$, and $\bar{w} = |S'|/|S|$, the fitness of $S$. Then

$$\bar{w} q'_i = \frac{|S'_i|}{|S_i|} \cdot \frac{|S'_i|}{|S|} = q_i W_i.$$ 

Sum this: $\bar{w} = \bar{W}_i$. Use these to reformulate the right side of Price’s Eq. (1):

$$\bar{w} \sum \Delta q_i Z_i = \bar{w} \sum q'_i Z_i - \bar{w} \sum q_i Z_i = \sum q_i W_i Z_i - \bar{W}_i \bar{Z}_i = \text{Cov}(W_i, Z_i).$$

$$\bar{w} \sum q'_i \Delta Z_i = \sum q_i W_i \Delta Z_i = \bar{W}_i \Delta Z_i.$$ 

Substitute in Eq. (1):

$$\bar{w} \Delta \tau = \text{Cov}(W_i, Z_i) + \bar{W}_i \Delta Z_i.$$  \hspace{1cm} (2)

This is the covariance form of Price’s equation. The selection $\sum \Delta q_i Z_i$ in Eq. (1) is $\text{Cov}(W_i, Z_i)/\bar{w}$ in Eq. (2). The better the $Z_i$ covary with the fitnesses $W_i$, the stronger the selection for it. Intuitive!

We will use Eq. (2) to derive the criterion for kin selection (better called inclusive fitness), Eq. (7), and the criterion for group selection (better called multilevel selection), Eq. (10).

---

2 One can also see directly that $q'_i = q_i |W_i|/\bar{w}$.  
3 A straightforward calculation shows that $\bar{w} \bar{Z}_i Z_i - \bar{W}_i \bar{Z}_i = (\bar{W}_i - \bar{W}_i)(\bar{Z}_i - \bar{Z}_i)$, just as with unweighted averages. The right side shows the meaning of covariance.
A Biological Interpretation

$S$ is a set of individuals, each with a value $z$ of some phenotypic trait.

(Natural selection acts on phenotypes, regardless of their genetic basis.)

$S_i$ is the set of individuals with value $z_i$ for the trait.

$S'$ is the set of offspring of the individuals, each with a value $z'$ of the trait.

$S'_i$ is the set of offspring of individuals in $S_i$.

A special case of Eq. (2) occurs when each set $S_i$ is an individual $j$ in $S$. At this individual level biologists usually set the transmission bias to zero:

$$\overline{w} \Delta z = \text{Cov}(w_j, z_j).$$  \hspace{1cm} (3)

Here $w_j$ is the fitness of individual $j$ – the number of its offspring – and $z_j$ is the $z$-value of individual $j$. The covariance represents the natural selection part of evolution.

Both kin and group selection use the notion of a group. “A group is defined as a set of individuals that influence each other’s fitness with respect to a certain trait but not the fitness of those outside the group.” Geographically proximate individuals need not be a group, even if they interact socially. Groups will be represented by the sets $S_i$ in our formalism.

---

4 For asexual reproduction. For sexual reproduction, each individual contributes $\frac{1}{2}$ of its offspring to its $S'_i$.

5 For asexual reproduction simply assume that there are no mutations. Then $\Delta z_j = 0$, so the transmission bias is zero. This doesn’t work for sexual reproduction because we expect $\Delta z_j \neq 0$ even with no mutations. According to Price, “If meiosis and fertilization are random with respect to [an altruistic gene], the [transmission bias] will be zero except for statistical sampling effects (random drift), and these will tend to average out to give [Eq. (3)].” G. Price, *Selection and Covariance*, Nature 227, 520 (1970). I don’t follow Price very well here.

Also: “[Eq. (3)] . . . concerns only selection itself, not the response to selection; the latter depends on the fidelity of transmission across generations, hence on facts about the inheritance mechanism etc. Throughout this paper, our concern will be with selection itself, hence the within-generation change in $z$, rather than the response to selection. So we do not need to worry about the genetic basis of $z$, whether it is heritable, and so on. This separation of selection from the response to selection is standard fare in quantitative genetics.” S. Okasha, *Multi-Level Selection, Price’s Equation and Causality*, Preprint (July 2003).

www.lse.ac.uk/CPNSS/pdf/DP_withCover_Causality/CTR13-03-C.pdf

6 For asexual reproduction. For sexual reproduction divide by 2.

7 So $W_i$ is the $w$-average in $S_i$ and $\overline{w}$ is the $w$-average in $S$.

Kin Selection

Many organisms exhibit altruistic behavior which benefits others’ fitness at a potential cost to their own. For example, some animals give an alarm call to their neighbors when they see a predator, endangering themselves. This poses a serious challenge to the theory of natural selection, if understood as “survival of the fittest”. Darwin was well aware of the problem.

The theory of kin selection points out that “survival of the fittest” means “survival of the fittest genes”, “not survival of the fittest individuals”. The theory is due to William Hamilton. He gave two, very different, versions.

Here is a caricature of the first (1964) version \(^9\) to help gain intuition. It omits important points. Suppose that an altruistic act confers a fitness benefit \(b > 0\) on a beneficiary at a fitness cost \(c > 0\) to the altruist. Let \(r\) be the probability that they share any particular gene, in particular a gene for altruism. Thus \(r = 1\) for identical twins, \(\frac{1}{2}\) for siblings, \(\frac{1}{8}\) for cousins, . . . .

Then Hamilton’s rule states that altruistic acts will increase the prevalence of the altruistic gene if and only if \(rb > c\), i.e., the probable benefit to the gene exceeds the probable cost. Intuitive! It is one way to see that a trait can be favored despite being unfavorable to its carriers.

Hamilton: “An animal acting on this principle would sacrifice its life if it could thereby save more than two brothers” \(^10\) (e.g., \(b = 3c\), so \(rb = \frac{1}{2}(3c) > c\)).

Hamilton’s second (1975) version is based on Price’s equation. One possibility for assigning \(z\) values to an altruistic trait is to set \(z = 1\) if an individual possesses the trait and \(z = 0\) if not.

Individual \(j\) belongs to some group \(S_i\). Let \(y_j\) be the \(z\)-average in \(S_i\). Write a least squares approximation to \(w_j\), including the residual \(\epsilon_j\):

\[
w_j = w_0 + by_j - cz_j + \epsilon_j.
\] (4)

Take the covariance with \(z_j\):

\[
\text{Cov}(w_j, z_j) = \text{Cov}(w_0, z_j) + b\text{Cov}(y_j, z_j) - c\text{Cov}(z_j, z_j) + \text{Cov}(\epsilon_j, z_j).
\] (5)

In this equation \(\text{Cov}(w_0, z_j) = 0\), as \(w_0\) is constant; \(\text{Cov}(z_j, z_j) = \text{Var}(z_j)\); and \(\text{Cov}(\epsilon_j, z_j) = 0\). \(^{11}\) Define \(r = \text{Cov}(y_j, z_j)/\text{Var}(z_j)\). Substitute into Eq. (3):

\[
\bar{w} \Delta z = (rb - c)\text{Var}(z_j).
\] (6)

In particular, we have Hamilton’s rule (assume \(\bar{w} \neq 0, \text{Var}(z_j) \neq 0\),

\[
\Delta z > 0 \iff rb > c.
\] (7)

---


\(^{11}\)See p. 14 for \(\text{Cov}(\epsilon_j, z_j) = 0\).
The $r = \text{Cov}(y_j, z_j) / \text{Var}(z_j)$ here is more general than the relatedness $r$ in the first version of Hamilton’s rule. The covariance shows that $r$ is “A statistical tendency for the recipients of altruism to be altruists themselves.” Hamilton: “Kinship should be considered just one way of getting positive regression [i.e., $r$] of genotype in the recipient, and it is this positive regression that is vitally necessary for altruism.”

The term *kin selection*, though widely used, does not accurately describe this second version of Hamilton’s rule. *Inclusive fitness* is better and often used. Nevertheless, we will continue to use *kin selection*.

Even $r < 0$ (i.e., $\text{Cov}(y_j, z_j) < 0$) is possible, making possible the phenomenon of *spite*. By definition, a spiteful trait has $b < 0$ as well, so $rb > 0$. This makes $rb > c$ possible. Then an individual possessing the trait harms itself and is harmed by others in its group. Yet such traits can be favorably selected!

J. Birch published a useful comparison of the two versions of Hamilton’s rule.

Most discussions of kin selection have defined $b$ in terms of the fitness benefit bestowed by individuals on their group. Equation (4) is different: it defines $b$ in terms of the fitness benefit bestowed on individuals by their group. This is the *direct fitness* approach to kin selection. “Direct fitness can be mathematically easier to work with and has recently emerged as the preferred approach of theoreticians.” And “These two approaches are computationally equivalent.”

**Kin Selection Acceptance**

Kin selection is widely accepted. Yet it is contentious. We cite several opinions.

There are cases where “a striking quantitative fit between the predictions of kin selection theory and the empirical patterns that are observed.” Several studies are cited.

“Laboratory experiments show that [salamander] cannibals preferentially consume less related individuals. We established that the necessary condition for kin selection, Hamilton’s rule, was met.”

---

16 P. D. Taylor et al, *Direct fitness or inclusive fitness: how shall we model kin selection?* Journal of Evolutionary Biology **20**, 301 (2007). (But see minor caveats about equivalence.)
“While many studies have provided qualitative support for Hamilton’s rule, quantitative tests have not yet been possible due to the difficulty of quantifying the costs and benefits of helping acts. Here we use a simulated system of foraging robots . . . over hundreds of generations of selection in populations with different $c/b$ ratios, [and] show that Hamilton’s rule always accurately predicts the minimum relatedness necessary for altruism to evolve.”

A criticism of this study: “Their study is based on computer simulations and is not an empirical test. No actual robots were used to generate the results concerning Hamilton’s rule. Moreover, the computer simulation is arranged such that altruistic behavior spreads if $br > c$. No other outcome would have been possible. It is not surprising that Hamilton’s rule holds in a computer simulation that is specifically designed to validate it.”

Gardner et al. provide a listing of and rebuttals to claims that kin selection is limited in its application.

A pair of titles: What’s wrong with inclusive fitness? and There is nothing wrong with inclusive fitness.

A 2010 paper in Nature by M. Nowak, C. Tarnita, and E. O. Wilson claims that “kin selection almost never holds”. The paper caused a furor. A rebuttal with 137 (!) signatories, followed by a reply by the original authors, was also printed in Nature. The original authors have also posted a “brief statement”.

In 2013 Nowak and Wilson continued their attack with Limitations of inclusive fitness, this time with B. Allen. From the paper: “The dominance of inclusive fitness theory has held up progress in this area for many decades.” And “There is no problem in evolutionary biology that requires an analysis based on inclusive fitness.”

In 2015, David Queller and colleagues claimed that the conclusions of the 2010 paper are wrong, and in fact their “modeling strategy yields results that confirm important insights from kin selection and inclusive fitness.”

The paper by Birch is a commentary on the debate. He asserts that part of the problem is that Nowak et al. have in mind the kin selection version of Hamilton’s rule and their opponents have in mind the inclusive fitness version.

---

21A. Gardner et al., The genetical theory of kin selection, ibid., Table 2.
25A brief statement about inclusive fitness and eusociality, ibid.
26PNAS 110 20135 (2013).
28Ibid.
Group Selection

The kin selection explanation of altruism points out that altruists can share genes with others in their group. If \( rb > c \), then the altruistic trait will be favorably selected.

The theory of group selection points out that altruists in a group can make the group fitter. We will show that under suitable conditions the altruistic trait will then be favorably selected.

Recall Price’s Eq. (2) for the groups \( S_i \):

\[
\pi \Delta z = \text{Cov}(W_i, Z_i) + W_i \Delta Z_i.
\]

Call the selection \( \text{Cov}(W_i, Z_i) \) the between-group selection.

Now apply Eq. (3) to \( S_i \) (rather than \( S \)): \( W_i \Delta Z_i = \text{Cov}_i(w_j, z_j) \). Take the weighted average over \( i \): \( W_i \Delta Z_i = \text{Cov}_i(w_j, z_j) \). Call this selection average \( \text{Cov}_i(w_j, z_j) \) the within-group selection.

Substitute in Eq. (8):

\[
\pi \Delta z = \text{Cov}(W_i, Z_i) + \text{Cov}_i(w_j, z_j).
\]

Thus (assume \( \pi \neq 0 \))

\[
\Delta z > 0 \iff \text{Cov}(W_i, Z_i) + \text{Cov}_i(w_j, z_j) > 0.
\]

In group selection terms, a trait \( z \) is altruistic if between-group selection is positive (\( \text{Cov}(W_i, Z_i) > 0 \)), “good for groups”, and within-group selection is negative (\( \text{Cov}_i(w_j, z_j) < 0 \), “bad for individuals”. Equation (10) shows that the trait is favored (\( \Delta z > 0 \)) if and only if the between-group selection overpowers the within-group selection. Intuitive! It is one way to see that a trait can be favored despite being unfavorable to its carriers.

Kin selection and group selection divvy up the same \( \pi \Delta z \) on the left sides of Eqs. (6) and (9) in different ways on their right sides. For kin selection, the right side comes from applying Price’s equation to individuals \( j \) in \( S \). For group selection, it comes from applying the equation to the groups \( S_i \).\(^{29}\)

\(^{29}\)Both right sides use the approximation Eq. (3) at the individual level, i.e., they set the transmission bias at the individual level to zero.

Old vs. New Group Selection

Group selection was the generally accepted explanation of altruism into the 1960’s, culminating with the publication of V.C. Wynne-Edwards’ book *Animal Dispersion in Relation to Social Behavior* (1962). The mathematics in the book is limited to graphs.

Hamilton’s kin selection (1964) provided an alternative mathematical explanation. J. Maynard Smith’s *Group selection and kin selection* (1964)\(^{31}\) and George Williams’ *Adaptation and Natural Selection* (1966) both advocated kin selection and rejected group selection. Richard Dawkins’ semi-popular *The Selfish Gene* (1976) followed in the same vein. They were such powerful voices that mention of group selection became pretty much forbidden.

Except for Hamilton, who accepted group selection. He writes (1998) of “telling [Price] enthusiastically [in 1974] that through a ‘group level’ extension of his formula I now had a far better understanding of group selection and was possessed of a far better tool for all forms of selection acting at one level or at many than I had ever had before.”\(^{32}\)

The “new” group selection, based on Price’s equation as above, is different from the “old” group selection of Wynne-Edwards.

“One way of conceptualizing the difference between the old and new group selection models is that the new group selection models rely on within population group selection, whereas old group selection theory worked on between population group selection.”\(^{33}\)

“Most formal [new] group selection models . . . aim to explain the evolution of an individual character, often altruism, in a group-structured population.”\(^{34}\) The emphasis (in the original) on individual character is opposed to group character. The quote exactly describes Eq. (10), with \(\Delta z\) over individuals on the left, and group quantities on the right.

“Everyone agrees that group selection occurs.”\(^{35}\) “Unfortunately”, O. Eldakar and D. S. Wilson add, “‘everyone’ refers only to those who seriously study the subject.”\(^{36}\) (I’ll come to Dawkins later.)

Group Selection Examples

George Williams believed that group selection is rare. He did accept one example in his 1966 book: the “only convincing evidence for the operation of group selection” was a study on house mice.\(^{37}\) Then from *The Dawn of*
Darwinian Medicine (1991): “The evolutionary outcome will depend on relative strengths of within-host and between-host competition in pathogen evolution. This is a clear example of group vs. individual (clone) selection for altruism.”

Many thousands of species have female biased sex ratios. Colwell (1981) explained unbiased and biased sex ratios in group selection terms. Williams (1992): “I think it is desirable …to realize that selection in female-biased Mendelian populations favors males, and that it is only the selection among such groups that can favor the female bias.” According to Avilés (2012), “Although the initial reaction to Colwell’s interpretation was rather negative, his views are now generally accepted among those close to the field of sex ratios.”


Chuang et al. (2009) give a vivid description of laboratory group selection experiments with controlled parameters.

Some titles: Multilevel selection in natural populations of impatiens capensis; Intrademic Group Selection Imposed by a Parasitoid-Host Interaction; Reproductive Conflicts and Levels of Selection in the Ant Pristomyrmex pungens; Density-Dependent Multilevel Selection in the Great Lakes Sea Rocket; The Influence of Environmental Variation on Group and Individual Selection in a Cress; Group selection on the boundary of a stable population.

From a 2014 paper about social spiders: “Our study shows group selection acting in a natural setting, on a trait known to be heritable, and that has led to a colony-level adaptation.”

The opinion (first?) expressed by Williams, that group selection requires special conditions and so must be rare, is widespread. However, simulations have shown that “group selection is able to override counteracting individual selection under a much broader range of conditions than previously believed” and that “under general assumptions and a wide range of parameter values, predatory restraint may evolve as a group adaptation.”

---

40 G. Williams, Natural Selection: Domains, Levels and Challenges, p. 49.
43 J. S. Chuang et al., Simpson’s Paradox in a Synthetic Microbial System, Science 323, 272 (2009). The term “group selection” is not used, but see the group selection criterion Eq. (10) at the end of the caption for Figure 1. The figure is quite instructive.
Kin and Group Selection are Equivalent

Starting with \( z \) values on partitioned sets \( S \) and \( S' \) we defined \( \Delta \bar{z} = \bar{z}' - \bar{z} \). Then using the biological interpretation on Page 3 we derived two conditions, each equivalent to \( \Delta \bar{z} > 0 \): the condition for kin selection in Eq. (7) and the condition for group selection in Eq. (10). Conclusion: kin selection and group selection are equivalent. If you accept one, then you must accept of the other. Don’t just take my word calculations for it:

“There is widespread agreement that group selection and kin selection … are formally equivalent.”

“The mathematical equivalence of [inclusive fitness and group selection theories] has long been known.”

“The inclusive-fitness and group-selection models presented here do not represent alternative processes. Nor is one a subset of the other. Instead, they are simply alternative ways of viewing the same selection process.”

“It has long been understood that the kin selection and multilevel [group] selection approaches to social evolution are mathematically equivalent, and merely represent different partitions of the same evolutionary process.”

“Kin selection and group selection are now broadly understood to describe the same evolutionary process from complementary perspectives.”

A choice between kin and group selection depends on technical convenience and insight obtained, which might differ from application to application. In some applications it might be that the mathematics of kin selection is best but the understanding from group selection is best. Opinions:

“At one level, kin selection and group selection are just different ways of doing the maths or conceptualizing the evolutionary process. However, from a practical point of view, it could not be clearer that the kin selection approach is the more broadly applicable tool that we can use to understand the natural world.” This is the prevailing view.

With a kin selection explanation, Eq. (7), “the idea that altruism is good for the group but bad for the individual has been lost. … What is good for the individual can conflict with what is good for the group [Eq. (9)]. The concept of adaptation should reflect this fact.”

Group selection “provides valuable insights into the evolution of social behavior that are not so easy to see when selection is viewed as an individual-level or gene-level process.”

---

48 Marshall, ibid.
50 A. Gardner et al., The genetical theory of kin selection, ibid.
51 A. Wild, ibid.
Coyne, Pinker, and Dawkins

They deserve no space here, as they have contributed nothing but confusion about group selection. But they are outspoken and heard, so I address them. Coyne, *The demise of group selection*,\(^{55}\) and Pinker, *The False Allure of Group Selection*,\(^{56}\) attack group selection, with Dawkins cheering them on.\(^{57,58}\) They mostly beat the dead horse of old group selection.

Coyne, first sentences: “The idea that adaptations in organisms result from ‘group selection’ (selection among groups that differentially bud off subgroups, with those having good ‘group traits’ becoming more numerous), . . . has undergone a bit of resurgence in popular culture. This is in stark contrast to the views of most evolutionary biologists.” I don’t see subgroups budding off in Eq. (9). And “most evolutionary biologists” doesn’t include several I’ve quoted.

I quoted a paper endorsing the equivalence of kin and group selection, while finding kin selection “more broadly applicable”.\(^{59}\) Coyne describes the paper as “three smart biologists taking group selection apart”. He suggests reading the paper starting at a specific paragraph. In it I find “We emphasized that the new group selection methodology is not only correct and a potentially useful tool”. I don’t read this as “taking group selection apart”. At least Coyne keeps the discussion at a high level, writing of “the intellectual vacuity of group selection.”

In February 2013 Coyne noticed a 2009 textbook claim that 1-1 sex ratios are evidence against group selection. He posted *Another case of individual selection trumping group selection* in his blog.\(^{60}\) He cites only Fisher’s (1930) argument for a 1-1 ratio, mentioning only in passing “some exceptions”. But what about the exceptions? Coyne does not refer to the group selection explanation of unbiased and biased sex ratios cited above.

Pinker’s essay is mostly about the application of kin and group selection to human behavior. These notes are about the mathematical foundation of kin and group selection, so I only address Pinker’s comments about this.

Pinker writes: “The more carefully you think about group selection, the less sense it makes.” He means the old group selection: “This brings us to the familiar problem which led most evolutionary biologists to reject the idea of group selection in the 1960s.” He also writes of “groups budding off new groups”. I still don’t see this in Eq. (9).

\(^{55}\)https://whyevolutionistrue.wordpress.com/2012/06/24/the-demise-of-group-selection/.

\(^{56}\)http://edge.org/conversation/the-false-allure-of-group-selection.


\(^{58}\)http://www.richarddawkins.net/news_articles/2012/7/13/pinker-s-reply-on-group-selection.

\(^{59}\)S. West, A. Griffen, A. Gardner, ibid.

\(^{60}\)https://whyevolutionistrue.wordpress.com/2013/02/27/another-case-of-individual-selection-trumping-group-selection.
Twenty-three comments follow the essay. David Queller, among others, defends group selection: “Modern group selection theory is as mathematically rigorous as individual selection or inclusive fitness theory. I say this despite being someone who favors the inclusive fitness approach and whose entire career has been based on it.”

Pinker then replies. At the start he writes “The theory is mathematically equivalent to standard evolutionary theories based on gene selection or inclusive fitness, so the two theories make identical predictions and can never be empirically distinguished.” He does not say that this is not the old group selection attacked in his essay. The discussion of “this new theory of group selection” begins about half way through the reply. What follows is too diffuse for me to summarize. Read it yourself. He seems to forget that the new group selection and kin selection “can never be empirically distinguished.”

Dawkins deserves credit for helping to abolish the old group selection with his spectacularly influential The Selfish Gene (1976) and also The Extended Phenotype (1982). He deserves blame for helping to suppress the new group selection with his across-the-board rejection of all group selection.

Dawkins (1982): “The years [after] Darwin [saw] an astonishing retreat . . . , a lapse into sloppily unconscious group-selectionism. . . . It is only in recent years . . . that the stampede has been halted and and turned. We painfully struggled back, harassed by sniping from a Jesuitically sophisticated dedicated neo-group-selectionist rearguard, until we finally regained Darwin’s ground.” Whew!

The word “stampede” is from Hamilton (1975), who Dawkins’ quotes in support of his views: “Almost the whole field stampeded in the [group selection] direction where Darwin had gone circumspectly or not at all.” But Dawkins gets no support from Hamilton here. For Hamilton is writing about the old group selection. On the very next page we find “A recent reformulation of natural selection [by Price] can be adapted to show how two successive levels of the subdivision of a population contribute separately to the overall natural selection.” He then derives the group selection Eq. (9).

Dawkins has not let up. In 2012 he wrote “Group selection: the poorly defined and incoherent view . . . .” D. S. Wilson has strong criticism.

64 Hamilton, Innate social aptitudes of man, ibid., p. 331
Odds and Ends

Students must be disabused of wrongheaded group selection beliefs, which a biologist friend tells me are common. But it is wrong to tell them that group selection explanations are always wrong.

We have suffered two generations of a scientific scandal: everyone “knows” that group selection is bad science, making it impossible to benefit from its insights. How long will this wrongheaded meme persist?

“The range of circumstances in which inclusive fitness is known to apply is always being extended. . . . That range is still quite small, and there is a long way to go to cover the situations that most empirical biologists would consider usual.”

If \( z = w \) in Eq. (3), i.e., the trait is fitness, then the equation becomes \( \overline{w}\Delta\overline{w} = \text{Var}(w) \); for a given \( \overline{w} \) the change in fitness due to natural selection is proportional to its variance. This is Fisher’s fundamental theorem of natural selection. Fitness increases!

As is well known, the quantity \( r = \text{Cov}(y_j, z_j)/\text{Var}(z_j) \) in Eq. (6) is the slope of the least squares line \( y = r z + s \) for the points \((z_j, y_j)\). For the line is given by

\[
\begin{bmatrix}
\sum z_j^2 & \sum z_j \\
\sum z_j & n
\end{bmatrix}
\begin{bmatrix}
r \\
s
\end{bmatrix}
= 
\begin{bmatrix}
\sum z_j y_j \\
\sum y_j
\end{bmatrix}.
\]

Solving by Cramer’s rule gives
\( r = \text{Cov}(z_j, y_j)/\text{Var}(z_j) \).

George Price led a fascinating life, which ended in suicide in 1975. For a biographical sketch see Gardner.


---

67 Recall that the transmission bias is not included in Eq. (3).
70 [http://linguafranca.mirror.theinfo.org/0007/altruist.html](http://linguafranca.mirror.theinfo.org/0007/altruist.html)
Least Squares

This page justifies the claim on Page 4 that $\text{Cov}(\epsilon_j, z_j) = 0$.

Least squares problems can be formulated using the matrix equation

$$
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix}
= 
\begin{bmatrix}
  1 & x_{11} & x_{21} & \cdots & x_{m1} \\
  1 & x_{12} & x_{22} & \cdots & x_{m2} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  1 & x_{1n} & x_{2n} & \cdots & x_{mn}
\end{bmatrix}
\begin{bmatrix}
  \beta_0 \\
  \beta_1 \\
  \vdots \\
  \beta_m
\end{bmatrix}
+ 
\begin{bmatrix}
  \epsilon_1 \\
  \epsilon_2 \\
  \vdots \\
  \epsilon_n
\end{bmatrix}.
$$

Compactly: $y = X\beta + \epsilon$. The vector $y$ and matrix $X$ are given. A least squares solution is a vector $\beta$ which minimizes $\|\epsilon\|^2$.

Let $X'$ be the transpose of $X$. Compute:

$$
\|\epsilon\|^2 = \epsilon \cdot \epsilon = (y - X\beta) \cdot (y - X\beta) = y \cdot y - 2\beta \cdot X'y + \beta \cdot X'X\beta.
$$

Let $c$ be a constant vector and $C$ a constant matrix. Compute $\partial(\Sigma_i c_i \beta_i) / \partial \beta_k$ and $\partial(\Sigma_{ij} \beta_i C_{ij} \beta_j) / \partial \beta_k$ to show that

$$
\nabla(\beta \cdot c) = c \quad \text{and} \quad \nabla(\beta \cdot C\beta) = (C + C')\beta.
$$

Thus $\nabla\|\epsilon\|^2 = -2X'y + 2X'X\beta$. At a minimizing $\beta$, $\nabla\|\epsilon\|^2 = 0$. Then

$$
X'X\beta = X'y = X'(X\beta + \epsilon) = X'X\beta + X'\epsilon.
$$

Thus $X'\epsilon = 0$. The top entry of this vector equation says that $\overline{\epsilon} = 0$. Remembering this, the other entries say that $\text{Cov}((x_{i1}, \ldots, x_{in})', \epsilon) = 0$, $i = 1, \ldots, m$.

---

Alan Macdonald
macdonal@luther.edu
http://faculty.luther.edu/~macdonal

---

\textsuperscript{71}If $X$ is of full rank (which is $m + 1$ if $m + 1 \leq n$, which is usual), then $X'X$ is invertible. Thus there is a unique minimizing $\beta = (X'X)^{-1}X'y$. 

14