



# Price equation

In the theory of evolution and natural selection, the **Price equation** (also known as **Price's equation** or **Price's theorem**) describes how a trait or allele changes in frequency over time. The equation uses a covariance between a trait and fitness, to give a mathematical description of evolution and natural selection. It provides a way to understand the effects that gene transmission and natural selection have on the frequency of alleles within each new generation of a population. The Price equation was derived by George R. Price, working in London to re-derive W.D. Hamilton's work on kin selection. Examples of the Price equation have been constructed for various evolutionary cases. The Price equation also has applications in economics.<sup>[1]</sup>

The Price equation is a mathematical relationship between various statistical descriptors of population dynamics, rather than a physical or biological law, and as such is not subject to experimental verification. In simple terms, it is a mathematical statement of the expression "survival of the fittest".

## Statement

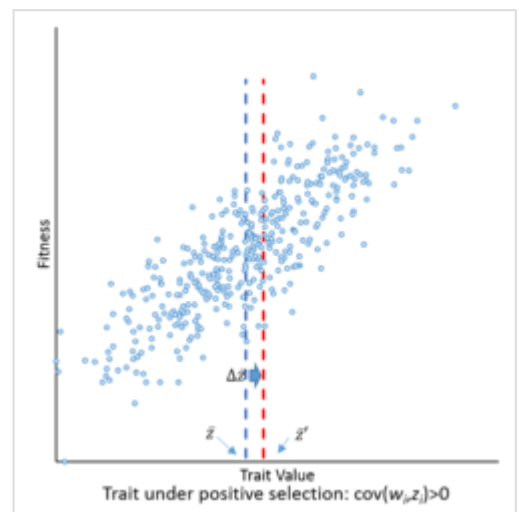
The Price equation shows that a change in the average amount  $z$  of a trait in a population from one generation to the next ( $\Delta z$ ) is determined by the covariance between the amounts  $z_i$  of the trait for subpopulation  $i$  and the fitnesses  $w_i$  of the subpopulations, together with the expected change in the amount of the trait value due to fitness, namely  $E(w_i \Delta z_i)$ :

$$\Delta z = \frac{1}{w} \text{cov}(w_i, z_i) + \frac{1}{w} E(w_i \Delta z_i).$$

Here  $w$  is the average fitness over the population, and  $E$  and  $\text{cov}$  represent the population mean and covariance respectively. 'Fitness'  $w$  is the ratio of the average number of offspring for the whole population per the number of adult individuals in the population, and  $w_i$  is that same ratio only for subpopulation  $i$ .

If the covariance between fitness ( $w_i$ ) and trait value ( $z_i$ ) is positive, the trait value is expected to rise on average across population  $i$ . If the covariance is negative, the characteristic is harmful, and its frequency is expected to drop.

The second term,  $E(w_i \Delta z_i)$ , represents the portion of  $\Delta z$  due to all factors other than direct selection which can affect trait evolution. This term can encompass genetic drift, mutation bias, or meiotic drive. Additionally, this term can encompass the effects of multi-level selection or group selection. Price (1972)



Example for a trait under positive selection

referred to this as the "environment change" term, and denoted both terms using partial derivative notation ( $\partial_{NS}$  and  $\partial_{EC}$ ). This concept of environment includes interspecies and ecological effects. Price describes this as follows:

Fisher adopted the somewhat unusual point of view of regarding dominance and epistasis as being environment effects. For example, he writes (1941): ‘A change in the proportion of any pair of genes itself constitutes a change in the environment in which individuals of the species find themselves.’ Hence he regarded the natural selection effect on  $\bar{M}$  as being limited to the additive or linear effects of changes in gene frequencies, while everything else – dominance, epistasis, population pressure, climate, and interactions with other species – he regarded as a matter of the environment.

—G.R. Price (1972), *Fisher's fundamental theorem made clear*<sup>[2]</sup>

## Proof

---

Suppose we are given four equal-length lists of real numbers<sup>[3]</sup>  $n_i, z_i, n'_i, z'_i$  from which we may define  $w_i = n'_i/n_i$ .  $n_i$  and  $z_i$  will be called the parent population numbers and characteristics associated with each index  $i$ . Likewise  $n'_i$  and  $z'_i$  will be called the child population numbers and characteristics, and  $w_i$  will be called the fitness associated with index  $i$ . (Equivalently, we could have been given  $n_i, z_i, w_i, z'_i$  with  $n'_i = w_i n_i$ .) Define the parent and child population totals:

$$n \stackrel{\text{def}}{=} \sum_i n_i \quad n' \stackrel{\text{def}}{=} \sum_i n'_i$$

and the probabilities (or frequencies):<sup>[4]</sup>

$$q_i \stackrel{\text{def}}{=} n_i/n \quad q'_i \stackrel{\text{def}}{=} n'_i/n'$$

Note that these are of the form of probability mass functions in that  $\sum_i q_i = \sum_i q'_i = 1$  and are in fact the probabilities that a random individual drawn from the parent or child population has a characteristic  $z_i$ . Define the fitnesses:

$$w_i \stackrel{\text{def}}{=} n'_i/n_i$$

The average of any list  $x_i$  is given by:

$$E(x_i) = \sum_i q_i x_i$$

so the average characteristics are defined as:

$$z \stackrel{\text{def}}{=} \sum_i q_i z_i \quad z' \stackrel{\text{def}}{=} \sum_i q'_i z'_i$$

and the average fitness is:

$$w \stackrel{\text{def}}{=} \sum_i q_i w_i$$

A simple theorem can be proved:  $q_i w_i = \left(\frac{n_i}{n}\right) \left(\frac{n'_i}{n_i}\right) = \left(\frac{n'_i}{n'}\right) \left(\frac{n'}{n}\right) = q'_i \left(\frac{n'}{n}\right)$  so that:

$$w = \frac{n'}{n} \sum_i q'_i = \frac{n'}{n}$$

and

$$q_i w_i = w q'_i$$

The covariance of  $w_i$  and  $z_i$  is defined by:

$$\text{cov}(w_i, z_i) \stackrel{\text{def}}{=} E(w_i z_i) - E(w_i)E(z_i) = \sum_i q_i w_i z_i - wz$$

Defining  $\Delta z_i \stackrel{\text{def}}{=} z'_i - z_i$ , the expectation value of  $w_i \Delta z_i$  is

$$E(w_i \Delta z_i) = \sum_i q_i w_i (z'_i - z_i) = \sum_i q_i w_i z'_i - \sum_i q_i w_i z_i$$

The sum of the two terms is:

$$\text{cov}(w_i, z_i) + E(w_i \Delta z_i) = \sum_i q_i w_i z_i - wz + \sum_i q_i w_i z'_i - \sum_i q_i w_i z_i = \sum_i q_i w_i z'_i - wz$$

Using the above mentioned simple theorem, the sum becomes

$$\text{cov}(w_i, z_i) + E(w_i \Delta z_i) = w \sum_i q'_i z'_i - wz = wz' - wz = w \Delta z$$

where  $\Delta z \stackrel{\text{def}}{=} z' - z$ .

## Derivation of the continuous-time Price equation

Consider a set of groups with  $i = 1, \dots, n$  that are characterized by a particular trait, denoted by  $x_i$ . The number  $n_i$  of individuals belonging to group  $i$  experiences exponential growth:

$$\frac{dn_i}{dt} = f_i n_i$$

where  $f_i$  corresponds to the fitness of the group. We want to derive an equation describing the time-evolution of the expected value of the trait:

$$\mathbb{E}(x) = \sum_i p_i x_i \equiv \mu, \quad p_i = \frac{n_i}{\sum_i n_i}$$

Based on the chain rule, we may derive an ordinary differential equation:

$$\begin{aligned} \frac{d\mu}{dt} &= \sum_i \frac{\partial \mu}{\partial p_i} \frac{dp_i}{dt} + \sum_i \frac{\partial \mu}{\partial x_i} \frac{dx_i}{dt} \\ &= \sum_i x_i \frac{dp_i}{dt} + \sum_i p_i \frac{dx_i}{dt} \\ &= \sum_i x_i \frac{dp_i}{dt} + \mathbb{E}\left(\frac{dx}{dt}\right) \end{aligned}$$

A further application of the chain rule for  $dp_i/dt$  gives us:

$$\frac{dp_i}{dt} = \sum_j \frac{\partial p_i}{\partial n_j} \frac{dn_j}{dt}, \quad \frac{\partial p_i}{\partial n_j} = \begin{cases} -p_i/N, & i \neq j \\ (1 - p_i)/N, & i = j \end{cases}$$

Summing up the components gives us that:

$$\begin{aligned} \frac{dp_i}{dt} &= p_i \left( f_i - \sum_j p_j f_j \right) \\ &= p_i [f_i - \mathbb{E}(f)] \end{aligned}$$

which is also known as the replicator equation. Now, note that:

$$\begin{aligned} \sum_i x_i \frac{dp_i}{dt} &= \sum_i p_i x_i [f_i - \mathbb{E}(f)] \\ &= \mathbb{E}\{x_i [f_i - \mathbb{E}(f)]\} \\ &= \text{Cov}(x, f) \end{aligned}$$

Therefore, putting all of these components together, we arrive at the continuous-time Price equation:

$$\frac{d}{dt}\mathbb{E}(x) = \underbrace{\text{Cov}(x, f)}_{\text{Selection effect}} + \underbrace{\mathbb{E}(\dot{x})}_{\text{Dynamic effect}}$$

## Simple Price equation

---

When the characteristic values  $z_i$  do not change from the parent to the child generation, the second term in the Price equation becomes zero resulting in a simplified version of the Price equation:

$$w \Delta z = \text{cov}(w_i, z_i)$$

which can be restated as:

$$\Delta z = \text{cov}(v_i, z_i)$$

where  $v_i$  is the fractional fitness:  $v_i = w_i/w$ .

This simple Price equation can be proven using the definition in Equation (2) above. It makes this fundamental statement about evolution: "If a certain inheritable characteristic is correlated with an increase in fractional fitness, the average value of that characteristic in the child population will be increased over that in the parent population."

## Applications

The Price equation can describe any system that changes over time, but is most often applied in evolutionary biology. The evolution of sight provides an example of simple directional selection. The evolution of sickle cell anemia shows how a heterozygote advantage can affect trait evolution. The Price equation can also be applied to population context dependent traits such as the evolution of sex ratios. Additionally, the Price equation is flexible enough to model second order traits such as the evolution of mutability. The Price equation also provides an extension to Founder effect which shows change in population traits in different settlements

## Dynamical sufficiency and the simple Price equation

Sometimes the genetic model being used encodes enough information into the parameters used by the Price equation to allow the calculation of the parameters for all subsequent generations. This property is referred to as dynamical sufficiency. For simplicity, the following looks at dynamical sufficiency for the simple Price equation, but is also valid for the full Price equation.

Referring to the definition in Equation (2), the simple Price equation for the character  $z$  can be written:

$$w(z' - z) = \langle w_i z_i \rangle - wz$$

For the second generation:

$$w'(z'' - z') = \langle w'_i z'_i \rangle - w'z'$$

The simple Price equation for  $z$  only gives us the value of  $z'$  for the first generation, but does not give us the value of  $w'$  and  $\langle w_i z_i \rangle$ , which are needed to calculate  $z''$  for the second generation. The variables  $w_i$  and  $\langle w_i z_i \rangle$  can both be thought of as characteristics of the first generation, so the Price equation can be used to calculate them as well:

$$\begin{aligned} w(w' - w) &= \langle w_i^2 \rangle - w^2 \\ w(\langle w'_i z'_i \rangle - \langle w_i z_i \rangle) &= \langle w_i^2 z_i \rangle - w \langle w_i z_i \rangle \end{aligned}$$

The five 0-generation variables  $w$ ,  $z$ ,  $\langle w_i z_i \rangle$ ,  $\langle w_i^2 \rangle$ , and  $\langle w_i^2 z_i \rangle$  must be known before proceeding to calculate the three first generation variables  $w'$ ,  $z'$ , and  $\langle w'_i z'_i \rangle$ , which are needed to calculate  $z''$  for the second generation. It can be seen that in general the Price equation cannot be used to propagate forward in time unless there is a way of calculating the higher moments  $\langle w_i^n \rangle$  and  $\langle w_i^n z_i \rangle$  from the lower moments in a way that is independent of the generation. Dynamical sufficiency means that such equations can be found in the genetic model, allowing the Price equation to be used alone as a propagator of the dynamics of the model forward in time.

## Full Price equation

---

The simple Price equation was based on the assumption that the characters  $z_i$  do not change over one generation. If it is assumed that they do change, with  $z_i$  being the value of the character in the child population, then the full Price equation must be used. A change in character can come about in a number of ways. The following two examples illustrate two such possibilities, each of which introduces new insight into the Price equation.

### Genotype fitness

We focus on the idea of the fitness of the genotype. The index  $i$  indicates the genotype and the number of type  $i$  genotypes in the child population is:

$$n'_i = \sum_j w_{ji} n_j$$

which gives fitness:

$$w_i = \frac{n'_i}{n_i}$$

Since the individual mutability  $z_i$  does not change, the average mutabilities will be:

$$z = \frac{1}{n} \sum_i z_i n_i$$
$$z' = \frac{1}{n'} \sum_i z_i n'_i$$

with these definitions, the simple Price equation now applies.

### Lineage fitness

In this case we want to look at the idea that fitness is measured by the number of children an organism has, regardless of their genotype. Note that we now have two methods of grouping, by lineage, and by genotype. It is this complication that will introduce the need for the full Price equation. The number of children an  $i$ -type organism has is:

$$n'_i = n_i \sum_j w_{ij}$$

which gives fitness:

$$w_i = \frac{n'_i}{n_i} = \sum_j w_{ij}$$

We now have characters in the child population which are the average character of the  $i$ -th parent.

$$z'_j = \frac{\sum_i n_i z_i w_{ij}}{\sum_i n_i w_{ij}}$$

with global characters:

$$z = \frac{1}{n} \sum_i z_i n_i$$

$$z' = \frac{1}{n'} \sum_i z_i n'_i$$

with these definitions, the full Price equation now applies.

## Criticism

---

The use of the change in average characteristic ( $z' - z$ ) per generation as a measure of evolutionary progress is not always appropriate. There may be cases where the average remains unchanged (and the covariance between fitness and characteristic is zero) while evolution is nevertheless in progress. For example, if we have  $z_i = (1, 2, 3)$ ,  $n_i = (1, 1, 1)$ , and  $w_i = (1, 4, 1)$ , then for the child population,  $n'_i = (1, 4, 1)$  showing that the peak fitness at  $w_2 = 4$  is in fact fractionally increasing the population of individuals with  $z_i = 2$ . However, the average characteristics are  $z=2$  and  $z'=2$  so that  $\Delta z = 0$ . The covariance  $\text{cov}(z_i, w_i)$  is also zero. The simple Price equation is required here, and it yields  $0=0$ . In other words, it yields no information regarding the progress of evolution in this system.

A critical discussion of the use of the Price equation can be found in van Veelen (2005),<sup>[5]</sup> van Veelen *et al.* (2012),<sup>[6]</sup> and van Veelen (2020).<sup>[7]</sup> Frank (2012) discusses the criticism in van Veelen *et al.* (2012).<sup>[8]</sup>

## Cultural references

---

Price's equation features in the plot and title of the 2008 thriller film WAZ.

The Price equation also features in posters in the computer game BioShock 2, in which a consumer of a "Brain Boost" tonic is seen deriving the Price equation while simultaneously reading a book. The game is set in the 1950s, substantially before Price's work.

## See also

---

- The breeder's equation, which is a special case of the Price equation.

## References

---

1. Knudsen, Thorbjørn (2004). "General selection theory and economic evolution: The Price equation and the replicator/interactor distinction" (<http://ideas.repec.org/a/taf/jecmet/v11y2004i2p147-173.html>). *Journal of Economic Methodology*. **11** (2): 147–173. doi:10.1080/13501780410001694109 (<https://doi.org/10.1080%2F1350178041000169410>)

- 9). S2CID 154197796 (<https://api.semanticscholar.org/CorpusID:154197796>). Retrieved 2011-10-22.
2. Price, G.R. (1972). "Fisher's "fundamental theorem" made clear". *Annals of Human Genetics*. **36** (2): 129–140. doi:10.1111/j.1469-1809.1972.tb00764.x (<https://doi.org/10.1111/j.1469-1809.1972.tb00764.x>). PMID 4656569 (<https://pubmed.ncbi.nlm.nih.gov/4656569>). S2CID 20757537 (<https://api.semanticscholar.org/CorpusID:20757537>).
3. The lists may in fact be members of any field (i.e. a set on which addition, subtraction, multiplication, and division are defined and behave as the corresponding operations on rational and real numbers do
4. Frank, Steven A. (1995). "George Price's Contributions to Evolutionary Genetics" (<https://www.researchgate.net/publication/15724182>). *J. Theor. Biol.* **175** (3): 373–388. Bibcode:1995JThBi.175..373F (<https://ui.adsabs.harvard.edu/abs/1995JThBi.175..373F>). doi:10.1006/jtbi.1995.0148 (<https://doi.org/10.1006/jtbi.1995.0148>). PMID 7475081 (<https://pubmed.ncbi.nlm.nih.gov/7475081>). Retrieved Mar 19, 2023.
5. van Veelen, M. (December 2005). "On the use of the Price equation". *Journal of Theoretical Biology*. **237** (4): 412–426. Bibcode:2005JThBi.237..412V (<https://ui.adsabs.harvard.edu/abs/2005JThBi.237..412V>). doi:10.1016/j.jtbi.2005.04.026 (<https://doi.org/10.1016/j.jtbi.2005.04.026>). PMID 15953618 (<https://pubmed.ncbi.nlm.nih.gov/15953618>).
6. van Veelen, M.; García, J.; Sabelis, M.W.; Egas, M. (April 2012). "Group selection and inclusive fitness are not equivalent; the Price equation vs. models and statistics". *Journal of Theoretical Biology*. **299**: 64–80. Bibcode:2012JThBi.299...64V (<https://ui.adsabs.harvard.edu/abs/2012JThBi.299...64V>). doi:10.1016/j.jtbi.2011.07.025 (<https://doi.org/10.1016/j.jtbi.2011.07.025>). PMID 21839750 (<https://pubmed.ncbi.nlm.nih.gov/21839750>).
7. van Veelen, M. (March 2020). "The problem with the Price equation" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7133513>). *Philosophical Transactions of the Royal Society B*. **375** (1797): 1–13. doi:10.1098/rstb.2019.0355 (<https://doi.org/10.1098/rstb.2019.0355>). PMC 7133513 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7133513>). PMID 32146887 (<https://pubmed.ncbi.nlm.nih.gov/32146887>).
8. Frank, S.A. (2012). "Natural Selection IV: The Price equation" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3354028>). *Journal of Evolutionary Biology*. **25** (6): 1002–1019. arXiv:1204.1515 (<https://arxiv.org/abs/1204.1515>). doi:10.1111/j.1420-9101.2012.02498.x (<https://doi.org/10.1111/j.1420-9101.2012.02498.x>). PMC 3354028 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3354028>). PMID 22487312 (<https://pubmed.ncbi.nlm.nih.gov/22487312>).

## Further reading

---

- Frank, S.A. (1995). "George Price's contributions to Evolutionary Genetics" (<http://stevefrank.org/reprints-pdf/95JTB-Price.pdf>) (PDF). *Journal of Theoretical Biology*. **175** (3): 373–388. Bibcode:1995JThBi.175..373F (<https://ui.adsabs.harvard.edu/abs/1995JThBi.175..373F>). CiteSeerX 10.1.1.136.7803 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.136.7803>). doi:10.1006/jtbi.1995.0148 (<https://doi.org/10.1006/jtbi.1995.0148>). PMID 7475081 (<https://pubmed.ncbi.nlm.nih.gov/7475081>).
- Frank, S.A. (1997). "The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis" (<http://stevefrank.org/reprints-pdf/97Evol-Causal.pdf>) (PDF). *Evolution*. **51** (6): 1712–1729. doi:10.2307/2410995 (<https://doi.org/10.2307/2410995>). JSTOR 2410995 (<https://www.jstor.org/stable/2410995>). PMID 28565119 (<https://pubmed.ncbi.nlm.nih.gov/28565119>).
- Gardner, A. (2008). "The Price equation" (<https://doi.org/10.1016/j.cub.2008.01.005>). *Curr. Biol.* **18** (5): R198 – R202. Bibcode:2008CBio...18.R198G (<https://ui.adsabs.harvard.edu/abs/2008CBio...18.R198G>). doi:10.1016/j.cub.2008.01.005 (<https://doi.org/10.1016/j.cub.2008.01.005>).



cub.2008.01.005). PMID 18334191 (<https://pubmed.ncbi.nlm.nih.gov/18334191>). S2CID 1169263 (<https://api.semanticscholar.org/CorpusID:1169263>).

- Grafen, A. (2000). "Developments of the Price equation and natural selection under uncertainty" (<http://users.ox.ac.uk/~grafen/cv/PriceUnc.pdf>) (PDF). *Proceedings of the Royal Society B*. **267** (1449): 1223–1227. doi:10.1098/rspb.2000.1131 (<https://doi.org/10.1098%2Frspb.2000.1131>). PMC 1690660 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1690660>). PMID 10902688 (<https://pubmed.ncbi.nlm.nih.gov/10902688>).
- Harman, Oren (2010). *The Price of Altruism: George Price and the search for the origins of kindness*. Bodley Head. ISBN 978-1-84792-062-1.
- Langdon, W.B. (1998). "8.1 Evolution of GP populations: Price's selection and covariance theorem". *Genetic Programming and Data Structures* (<https://archive.org/details/geneticprogrammingi00lang>). Springer. pp. 167 (<https://archive.org/details/geneticprogrammingi00lang/page/n183>)–208. ISBN 9780792381358.
- Price, G.R. (1970). "Selection and covariance" ([http://www0.cs.ucl.ac.uk/staff/ucacbb/ftp/papers/price\\_nature.pdf](http://www0.cs.ucl.ac.uk/staff/ucacbb/ftp/papers/price_nature.pdf)) (PDF). *Nature*. **227** (5257): 520–521. Bibcode:1970Natur.227..520P (<https://ui.adsabs.harvard.edu/abs/1970Natur.227..520P>). doi:10.1038/227520a0 (<https://doi.org/10.1038%2F227520a0>). PMID 5428476 (<https://pubmed.ncbi.nlm.nih.gov/5428476>). S2CID 4264723 (<https://api.semanticscholar.org/CorpusID:4264723>).
- Price, G.R. (1972). "Extension of covariance selection mathematics". *Annals of Human Genetics*. **35** (4): 485–490. doi:10.1111/j.1469-1809.1957.tb01874.x (<https://doi.org/10.1111%2Fj.1469-1809.1957.tb01874.x>). PMID 5073694 (<https://pubmed.ncbi.nlm.nih.gov/5073694>). S2CID 37828617 (<https://api.semanticscholar.org/CorpusID:37828617>).
- van Veelen, Matthijs; García, Julián; Sabelis, Maurice W. & Egas, Martijn (2010). "Call for a return to rigour in models" (<https://doi.org/10.1038%2F467661d>). Correspondence. *Nature*. **467** (7316): 661. Bibcode:2010Natur.467..661V (<https://ui.adsabs.harvard.edu/abs/2010Natur.467..661V>). doi:10.1038/467661d (<https://doi.org/10.1038%2F467661d>). PMID 20930826 (<https://pubmed.ncbi.nlm.nih.gov/20930826>).
- Day, T. (2006). "Insights from Price's equation into evolutionary epidemiology". *Disease Evolution*. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. Vol. 71. pp. 23–43. doi:10.1090/dimacs/071/02 (<https://doi.org/10.1090%2Fdimacs%2F071%2F02>). ISBN 9780821837535.
- "How to quit the Price equation: An online self-help tutorial" (<http://www.evolutionandgames.com/price>).
- "The Good Show" (<http://www.radiolab.org/2010/dec/14/>). *Radiolab*. Season 9. Episode 1. New York. 14 December 2011. WNYC.
- Markovitch; Witkowski; Virgo (2018). "Chemical heredity as group selection at the molecular level". arXiv:1802.08024 (<https://arxiv.org/abs/1802.08024>) [q-bio.PE (<https://arxiv.org/archive/q-bio/PE>)].