

van Veelen, Matthijs

**Working Paper**

## The generalized price equation

Tinbergen Institute Discussion Paper, No. TI 2024-034/I

**Provided in Cooperation with:**

Tinbergen Institute, Amsterdam and Rotterdam

*Suggested Citation:* van Veelen, Matthijs (2024) : The generalized price equation, Tinbergen Institute Discussion Paper, No. TI 2024-034/I, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/298041>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

TI 2024-034/I  
Tinbergen Institute Discussion Paper

# The Generalized Price Equation

*Matthijs van Veelen<sup>1</sup>*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900

# The Generalized Price Equation

**Abstract.** The main ingredient of this paper is the derivation of the Generalized Price equation. This generalizes the original Price equation in the sense that it produces a set of Price-like equations, one for every different underlying model that one could assume has generated the data. All of these different Price-like equations are identities, and all of them only have a meaningful interpretation if the data are indeed generated by the model they presuppose. The criteria for choosing between these different Price-like equations are the exact same as the criteria that standard statistics uses when choosing the right statistical model, based on the data. The original Price equation in regression form is the generalized Price equation that goes with the simplest linear model. The problem with the widespread misuse of the Price equation is caused by the fact that it loses its meaning if the data are not generated by this model – in the same way that any of the other Price-like equations lose their meaning if the data are not generated by the model they presuppose.

## 1. Introduction

Right from the beginning, it has been unclear if the Price equation [1] was intended as a tool for modeling, or if it was supposed to be applied to data. The terminology of regression coefficients suggests the latter, but the use in the population genetics literature mostly suggests the former. The paper in which the equation was presented is also inconsistent and goes back and forth between the two options (see Section 4.3 in [2] for a close reading of [1]). Here, we will first go with the second option, because that offers a way in which we can give a meaningful interpretation to the “regression coefficients” in what Price [1] called the “regression form” of his equation. Both the initial paper and the subsequent literature suggest there is a link between statistics and the regression form of the Price equation, without establishing what that link is. This paper does provide that link, and in doing so it formulates a generalized version of the Price equation that gives it the flexibility that a proper link with statistics requires. This Generalized Price equation also turns out to provide a solution for the problems with the Price equation identified in [2-5]. The insight that it provides also sheds light on exactly what the source of the controversy in the Price equation literature is.

In this paper, we will derive a general version of the Price equation. Besides the introduction of a bit of matrix notation, the first part of that derivation is totally standard, and leads to the Price equation as we know it, in what Price [1], [6] called the covariance form.

$$\bar{w}\Delta\bar{p} = \text{Cov}(w, p) + E(w\Delta p) \quad (\text{PE.C})$$

Then we combine this with a set of statistical models, all of which include 1) a constant term, and 2) a linear term for the set of genes, the selection of which is tracked by the variable  $\bar{p}$ , or the average  $p$ -score. For any model in this set, we find that in the covariance term, we can replace the realized fitnesses  $w_i$  with the estimated fitnesses  $\hat{w}_i$  according to the model. This leads to the generalized Price equation in covariance form, that, for any model in this set, reads

$$\bar{w}\Delta\bar{p} = \text{Cov}(\hat{w}, p) + E(w\Delta p) \quad (\text{GPE.C})$$

If we then unpack these covariances for any model in this set, we arrive at the generalized Price equation in regression form. We can for instance consider the set of models given by

$$w_i = \alpha + \sum_{r=1}^R \beta_r p_i^r + \varepsilon_i$$

The  $R$  is the largest exponent that is included in the polynomial of the model. Different choices for  $R$  therefore imply different models; for  $R = 1$  we get a linear model; for  $R = 2$  we get a quadratic model; for  $R = 3$  a cubic one, and so on. For every  $R$ , and thereby for every model in this set, we get a Price-like equation:

$$\bar{w}\Delta\bar{p} = \sum_{r=1}^R \hat{\beta}_r \text{Cov}(p, p^r) + E(w\Delta p) \quad (\text{GPE.R1})$$

The central point here is that there is not just one Price equation. There is a Price-like equation for every different model, as long as these models meet the minimal requirement of including a constant and a linear term. For rich enough settings (that is, for settings with sufficiently many different values that the  $p$ -scores can take) these Price-like equations moreover are generically proper different. That is to say, if we take any transition from a parent population to an offspring population, and combine this, for instance, with the Price equation for the linear model ( $R = 1$ ), with the Price equation for the quadratic model ( $R = 2$ ), then the  $\hat{\beta}_1$  in the Price equation for  $R = 1$  will almost always be different from the  $\hat{\beta}_1$  in the Price equation for  $R = 2$ , unless the  $p$ -score can only be 0 or 1.

The original Price equation in regression form is the generalized Price equation in regression form for the linear model. That implies that if the data are in fact generated by the standard linear model, writing *the* Price equation is a meaningful exercise. It also implies that if the data are not generated by the standard linear model, then writing *the* Price equation is not a meaningful exercise. For any set of data, one can write many Price-like equations, and all of them are identities. A meaningful Price equation exercise pertaining to data should really consist of using standard statistical tools for choosing between different statistical models. Any choice for a statistical model then automatically implies a choice for one of these Price-like equations in regression form. The confidence with which the data allow us to pick one of the models is then by definition equal to the confidence we can have that we picked the right Price-like equation to describe the population genetic dynamics.

The fact that there are more than one Price-like equation requires us to reflect on what the original Price equation is, and what we want to use it for. The Price equation is (i) an identity. That means that the left-hand side is equal to the right-hand side, whatever the parent population is and whatever the offspring population is. This is true for the original Price equation, and it is also true for all Price-like equations that we get by combining the Generalized Price equation with some model (choose some  $R$ ). The being an identity implies that it (ii) “always gets the direction of selection right”; if the left-hand side is negative, then so is the right-hand side. This is a somewhat riskier way of looking at the equation because it makes it tempting to see one side of the identity as an explanation of the other. Again, this property is shared by all Price-like equations. The main ingredient that I would like to add to this, and that the abundance of Price-like equations necessitates, is that we would like to

pick one of those equations on the basis of whether or not it accurately describes the underlying population genetical dynamics. If it does, then I use the word “meaningful”.

The main insight from the derivation of the generalized Price equation in regression form is therefore first and foremost positive and constructive, in the sense that it provides a recipe for finding an equation that is meaningful in the same way that the original Price equation is meaningful for the linear model. Besides that, it also helps us pinpoint what drives the misuse of the Price equation in the literature. Misapplying the Price equation comes down to applying the original Price equation (which is the generalized Price equation for the linear model) in a setting with a statistical model that is not linear, or that is general enough to allow for models other than the standard linear model. The misuse of the Price equation therefore is a special case of the chosen Price-like equation not matching the statistical model.

Besides this main result, this paper also serves as a guide to the Price equation. We will start simple, by introducing a haploid setting with asexual reproduction. The reason to start there is that it is the easiest. We then derive the normal, standard Price equation, and consider a simple example, where it tracks the absence or presence of a single gene. Then we move on to allow for a genetic measure that is not binary, and derive the general version of the Price equation, both in covariance form and, for a specific set of models, in regression form. This derivation helps establish that there is a variety of Price-like equations. It also shows that the answer to the question which one allows for a meaningful interpretation is given by standard statistical considerations involving sample size and model specification. All of this happens in Section 2.

In Section 3 we repeat this for diploid, sexually reproducing species. This is more interesting and allows us to illustrate the richness of possible underlying true models better. This section contains examples in which we assume heterozygote advantage, or dominance, and in which the genes we consider are sex-determining genes.

In Section 4 we further enrich the set of possible models, by allowing the model to not only include genes, the selection of which we track with the average  $p$ -score, but also take other genes into consideration, reflected in an average  $q$ -score. This introduces additional scope for misspecification.

In Section 5 we will describe how this carries over to the setting in which the Price equation is used for modelling.

In Section 6 we conclude, and reflect on a few related issues, including dynamic sufficiency.

In the twin TI discussion paper on the general version of Hamilton’s rule, we discuss how the use of the original Price equation, instead of its generalized version, has led to a longstanding disagreement concerning the generality of Hamilton’s rule, and how it can be resolved using the Generalized Price equation.

## 2. The Price equation for asexual reproduction

The original Price equation allows for any ploidy. Here we begin simple, by choosing a ploidy of 1. That keeps the introduction of the matrix notation as well as the first examples easy to follow.

We assume that the parent generation consist of  $n$  individuals, and the offspring generation consist of  $m$  individuals. We can represent who is whose offspring with an offspring matrix  $A$ . Let  $A_{ij} = 1$  if  $i$  is  $j$ 's parent, and  $A_{ij} = 0$  if not.

If we fix a member of the offspring generation, and sum over the parent generation, we get the number of parents per child. With asexual reproduction, all kids have one parent, which implies that this sum must be 1;  $\sum_{i=1}^n A_{ij} = 1$  for all offspring  $j$ . Any offspring matrix must have this property.

If we fix a member of the parent generation, and sum over the offspring generation, we get the number of children for that parent;  $a_i = \sum_{j=1}^m A_{ij}$  for all parents  $i$ . With asexual reproduction, this is also the fitness of the parent;  $w_i = a_i$ .

These two observations together have a straightforward implication. We can sum all elements of this matrix in two ways. The first is  $\sum_{j=1}^m (\sum_{i=1}^n A_{ij}) = \sum_{j=1}^m 1 = m$ . The second is  $\sum_{i=1}^n (\sum_{j=1}^m A_{ij}) = \sum_{i=1}^n w_i$ . Because these must be equal, the sum of all fitnesses must equal the number of individuals in the offspring generation;  $\sum_{i=1}^n w_i = m$ .

**Genes in the parent generation.** Suppose that we know the “dose” of a gene [1], or the  $p$ -score [7], for every individual in the parent population, and denote it by  $p_i$ ,  $i = 1, \dots, n$  for all parents. This dose may be restricted to be either 0 or 1, as it will in our first example, where this represents the presence or absence of a single gene. The idea of a dose, or a  $p$ -score, however, is that we also allow for this to be any measure for the presence or absence of a set of genes, and in particular genes that all contribute to a certain trait value. If we have two genes, both of which independently raise a certain trait value by the same amount, then a natural choice would be for the  $p$ -scores to be 0 if both are absent,  $\frac{1}{2}$  if one is present, and 1 if both are present. In general, this means that  $p_i$  can take values on some subset of the interval  $[0,1]$ .

**Which genes are passed on.** With sexual reproduction, there is randomness in which genes a parent passes on to their offspring. With asexual reproduction, offspring are just copies of their parents. That makes the matrix  $P$ , that represents the  $p$ -scores of the genes that are passed on, relatively easy. With asexual reproduction,  $P_{ij} = 0$  if individual  $i$  from the parent population is just not a parent of individual  $j$  in the offspring population (in other words  $P_{ij} = 0$  if  $A_{ij} = 0$ ). If individual  $i$  from the parent population is in fact the parent of individual  $j$  in the offspring population ( $A_{ij} = 1$ ), then  $P_{ij}$  is the  $p$ -score of the offspring. Here we assume that there is no mutation, which means that this is also the  $p$ -score of the parent. This implies a simple, straightforward relation between the  $p$ -scores in the offspring generation and the matrix  $P$ . If we fix an offspring, and sum over everyone in the parent generation, then all but one are not the parent of individual  $j$ , and for those,  $P_{ij} = 0$ . For the

actual parent, the matrix returns the  $p$ -score of the offspring, and hence  $\sum_{i=1}^n P_{ij} = p'_j$  for all offspring  $j$ .

**Change.** The average  $p$ -score in the parent generation is  $\bar{p} = \frac{\sum_{i=1}^n p_i}{n}$ . Some of the older papers on the Price equation do not use upper bars to indicate averages. Since it natural for  $p$  to represent the vector of  $p$ -scores in the parent generation,  $p = [p_1, \dots, p_n]$ , we do mostly add an upper bar to make sure there is no ambiguity regarding what this is:  $\bar{p}$  is the average of the  $p_i$ 's in the parent population. Similarly, the average  $p$ -score in the offspring generation is  $\bar{p}' = \frac{\sum_{j=1}^m p'_j}{m}$ .

**Derivation of the Price equation.** The key ingredient in the derivation of the Price equation is that one can calculate the sum of the  $p$ -scores by going over the individuals in the offspring generation in two ways. With sexual reproduction, this is a bit more complex, but with asexual reproduction, this is really quite simple.

1. One can go over all offspring, starting at offspring number 1 and ending at number  $m$ , and just add up their  $p$ -scores. With the relation between matrix  $P$  and vector  $p'$ , that amounts to  $\sum_{j=1}^m p'_j = \sum_{j=1}^m (\sum_{i=1}^n P_{ij})$ .
2. For every parent, we can add up the  $p$ -scores of their offspring. Then we can go over all parents, and add those numbers up. This amounts to a switch in summation order from  $\sum_{j=1}^m (\sum_{i=1}^n P_{ij})$  to  $\sum_{i=1}^n (\sum_{j=1}^m P_{ij})$ .

This switch allows us to write the average  $p$ -score in the offspring generation as

$$\bar{p}' = \frac{\sum_{j=1}^m p'_j}{m} = \frac{\sum_{j=1}^m \sum_{i=1}^n P_{ij}}{m} = \frac{\sum_{i=1}^n \sum_{j=1}^m P_{ij}}{m}$$

Then we can divide  $\sum_{j=1}^m P_{ij}$  by  $w_i$ , if we undo that by also multiplying by  $w_i$ , and we can subtract  $\sum_{i=1}^n w_i p_i$  if we undo that by also adding  $\sum_{i=1}^n w_i p_i$ . This way we get

$$\bar{p}' = \frac{\sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} \right)}{m} = \frac{\sum_{i=1}^n w_i p_i + \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right)}{m}$$

Then we add and subtract  $\frac{1}{n} \sum_{i=1}^n w_i \sum_{i=1}^n p_i$

$$\bar{p}' = \frac{\sum_{i=1}^n w_i p_i + \frac{1}{n} \sum_{i=1}^n w_i \sum_{i=1}^n p_i - \frac{1}{n} \sum_{i=1}^n w_i \sum_{i=1}^n p_i}{m} + \frac{\sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right)}{m}$$

Now if we look at the first term on the right-hand side, and then the second term in the numerator, we can use  $\sum_{i=1}^n w_i = m$  to simplify it. Because  $\frac{1}{n} \sum_{i=1}^n p_i$  is moreover the definition of the average  $p$ -score in the parent population, we can rewrite this whole equation as



$$\bar{p}' = \bar{p} + \frac{\sum_{i=1}^n w_i p_i - \frac{1}{n} \sum_{i=1}^n w_i \sum_{i=1}^n p_i}{m} + \frac{\sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right)}{m}$$

We can subtract  $\bar{p}$  on both sides and multiply this left and right by the average fitness  $\bar{w} = \frac{m}{n}$ . If we do that, we get

$$\bar{w}(\bar{p}' - \bar{p}) = \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i + \frac{1}{n} \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right)$$

A short way to write this is

$$\bar{w}\Delta\bar{p} = \text{Cov}(w, p) + E(w\Delta p) \quad (\text{PE.C})$$

where

$$\text{Cov}(w, p) = \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i$$

and

$$E(w\Delta p) = \frac{1}{n} \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right)$$

This is what Price [1], [6] calls the *covariance form*. Here it is important to realize that the term that is abbreviated as  $\text{Cov}(w, p)$  is *not* an actual covariance. I have made this point elsewhere<sup>1</sup> and while it seems nit-picking, it pays to be precise conceptually. A covariance is a property of a random variable, and that is not what the  $\text{Cov}(w, p)$ -term is here. If  $p_i$  and  $P_{ij}$  are data, as we assume they are here, then the term abbreviated as  $\text{Cov}(w, p)$  is the *sample* covariance. In Section 5, we switch to a modeling context. In that case,  $p_i$  represent a population state in a model, and the model is a statement about the probabilities of different transitions. Also here, the term abbreviated as  $\text{Cov}(w, p)$  is not a covariance; it is a summary statistic describing an aspect of the transition from one population state to another. Similarly, the second term on the right-hand side is *not* an expected value. If  $p_i$  and  $P_{ij}$  are data, it is the average difference between, on the one hand, the average  $p$ -score of the offspring of a parent and, on the other, its own  $p$ -score. Without mutation, this term is 0 for asexual reproduction by definition.

When we will derive the Generalized Price equation in Section 2.2, we will come back to this point in the derivation and proceed from there. In this section, on the other hand, we will simply follow Price [1], [6] in substituting  $\beta \cdot \text{Var}(p)$  for  $\text{Cov}(w, p)$ , where

<sup>1</sup> Page 414 of [2], and Box 1, page 66 of [5].

$$\text{Var}(p) = \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2$$

and

$$\beta = \frac{\text{Cov}(w, p)}{\text{Var}(p)}$$

If we do, then we get the following equation.

$$\bar{w}\Delta\bar{p} = \beta \cdot \text{Var}(p) + E(w\Delta p) \quad (\text{PE.R})$$

This is what Price [1], [6] calls *the regression form*.  $\text{Var}(p)$  is not really a variance, but if  $p_i$  and  $P_{ij}$  are data, it is the *sample* variance. The  $\beta$  is typically referred to as a regression coefficient, which it is if  $p_i$  and  $P_{ij}$  are data. We will get back to all of this below, but for now this is the final equation.

Our next step is to look at two specific examples. After that, we will return to the *covariance form* of the Price equation (PE.C) above, and derive the generalized version starting from there.

### Example 2.1

We start by assuming that the model captures one single gene – so  $p$ -scores are either 0 or 1 – and that this one gene has some (unknown) effect on fitness. In other words, we assume that the true model that describes how fitnesses depend on  $p$ -scores is

$$w_i = \alpha + \beta p_i + \varepsilon_i \quad (\text{A})$$

where  $\alpha$  is the baseline fitness,  $p_i = 0$  if the gene is absent,  $p_i = 1$  if it is present,  $\beta_1$  is the linear effect of the gene on fitness, and  $\varepsilon_i$  is a noise term.

We can, for example, take a parent population that consists of two individuals. The first has the gene and two offspring. The second does not have the gene and has no offspring. In other words,  $p = [1,0]$ ,  $w = [2,0]$  and

$$A = P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

The ingredients of the Price equation then are:  $\bar{w} = 1$ , as there are 2 parents and 2 offspring;  $\Delta\bar{p} = \frac{1}{2}$ , because the average  $p$ -score went from  $\frac{1}{2}$  to 1;  $E(w\Delta p) = 0$ , by definition with asexual reproduction without mutations;  $\text{Cov}(w, p) = \frac{1}{2}$ ,  $\text{Var}(p) = \frac{1}{4}$  and  $\beta = 2$ . The calculations are in Appendix A.

The Price equation therefore is:

$$\bar{w}\Delta\bar{p} = \beta \cdot \text{Var}(p) + E(w\Delta p)$$

$$1 \cdot \frac{1}{2} = 2 \cdot \frac{1}{4} + 0$$

In this case, the sample size is clearly not large enough for the  $\beta$  from the Price equation to estimate the true  $\beta$  from the model with any significance. The  $\beta$  in the Price equation therefore cannot be interpreted as the fitness effect of the gene.

At this point, there is some room for confusion, caused by the use of the word “population”. This is discussed on page 416 of [2], but it can be helpful to also elaborate a little on this here. Part of our intuition about statistics is shaped by polling or surveys. The setting for a poll or a survey is that there is a large population, of which we would like to know, for instance, what they would vote, if the election were held today; or how much time the average EU-citizen spends behind a screen; or what fraction of the people in the UK believes in evolution. The idea is then to randomly draw a sample from the population as a whole, and the larger the sample, and the more balanced it is, the larger the confidence that the average within the sample is close to the average of the population as a whole. On election day, we then find out the thing that the polls were trying to get at.

This shapes our intuition, in the sense that we are tempted to think that if the two individuals in our sample *are* the whole population (as they are in the example above with the Price equation), then there is no statistical uncertainty left, and we know everything there is to know. That would be the case, if the question at hand were “*how many offspring did the carriers of the gene get within this population, and how many did the non-carriers get?*” In our case, however, the question that we would be interested in, in order to be able to describe the population genetic dynamics, is “*how does the p-score affect fitnesses?*”. For answering the second question, having a parent population of size 2 will not be sufficient.

In order to illustrate the difference with a thought experiment, we can think of a relatively common versus an extremely rare genetic disorder. Suppose there is a genetic disorder that one in 1,000 people have. Assume also that we have data on everyone that was born with that genetic disorder in the UK in the years 2000 to 2005. In that case we would have a few thousand observations, which is more than enough to calculate the probability that people born with that genetic disorder survive into adulthood. If, within this thought experiment, we furthermore can assume that there are no changes in medical treatment, or other factors that could make being born in 2024 be different from being born in 2000, then for someone born with that genetic disorder in the UK today, the probability of surviving until the age of 18 is estimated with high accuracy by the realized survival rate within this population.

Now imagine another genetic disorder that only 1 in a million people are born with, and again assume that we have data for everyone that was born with this genetic disorder in the UK between 2000 and 2005. With the low incidence of this particular genetic disorder, this can easily leave us with less than a handful of observations only. Now even though in both cases we know the entire population within the bracket, for the extremely rare genetic disorder, we would not make predictions concerning the chances of survival until adulthood with the same confidence as we would for the much more common genetic disorder. For a doctor, the relevant question is what the probabilities are for her patient, and for this, knowing the whole population is very useful with the common genetic disorder, but much less so with the rare genetic disorder.

This serves to illustrate that, while with polling, the aim is to know the composition of the population, there are many settings in which that is not the underlying question. Statisticians therefore sometimes like to use the term “data generating process” instead of “population”. For polling, or market research, the question at hand is; if I were to pick a random member of the population, what would this person vote, or like to purchase. The data-generating process then would be that we ask a subset of people about their preferences. For an extremely rare genetic disease, however, the data generating process is that every once in a blue moon someone with that genetic disorder is born. This can render so few individuals that have the genetic disorder, that even knowing everyone in the population that does, will not give an answer to the question what for instance someone’s life expectancy is.

If we now go back to the Price equation, then with a small population, even if the Price equation encompasses the whole population, it may not always be right to interpret the  $\beta$  in it as the effect of the gene. It is a regression coefficient, and as such, it may be a noisy, and therefore unreliable, estimator of the true fitness effect of the gene, even if it is unbiased.

One can however also imagine a vector  $p$ , and matrices  $A$  and  $P$ , that together represent a much larger population, and therefore a much larger dataset. If this results in an estimate of the true  $\beta$  that is statistically significant (that is: the true  $\beta$  is most likely to be different from 0) then the  $\beta$  in the Price equation can be interpreted as the estimated effect of the gene in question on fitness. After the next example, from the derivation of the Generalized Price equation onwards, we will therefore use notation that differentiates between the true effect  $\beta$  and the estimator  $\hat{\beta}$  of the true effect.

All of this is under the assumption that model (A),  $w_i = \alpha + \beta p_i + \varepsilon_i$ , is accurately specified. In this example, we assumed that the  $p$ -score only reflects the absence or presence of one single gene, and therefore the  $p$ -score can only be 0 or 1. That implies that in this case, there is no room for underspecification, and this assumption does not need to be tested.

The main observation here is that the extent to which we can give a meaningful interpretation of the right-hand side of the Price equation is determined by standard statistical considerations. With small sample sizes, the  $\beta$  from the Price equation will, as an estimator of the true  $\beta$ , typically not be significant, even if the true  $\beta$  would really be different from 0. The low power of statistical testing then implies that there is really nothing we can say about the true  $\beta$ . Larger sample sizes allow us to infer more from the data, but even then, the interpretation is bound by the rules of statistics. If the  $\beta$  from the Price equation is statistically significant, it is still the (unbiased) *estimator* of the effect of the gene on fitness, and not the effect of the gene on fitness itself.

If the  $p$ -score only reflects the presence or absence of a single gene, and  $\beta$  is statistically significant, then the right-hand side of the Price equation has a meaningful interpretation; the  $\beta$  from the regression form of the Price equation is the estimated linear effect of this gene of fitness.

## Example 2.2

We now assume that the model captures two genes, both of which contribute to one and the same trait value. The possible  $p$ -scores therefore are 0 if both are absent;  $\frac{1}{2}$  if one is present; or 1 if both are present. If the  $p$ -scores translate to fitnesses in a linear way, then we are back in the setting of Example 2.1, just with one additional value that the  $p$ -score can take. We can however also allow for the possibility that the relation between  $p$ -scores and fitnesses is not linear. There are two steps between  $p$ -scores and fitnesses, so the non-linearity might arise in either of the two. It might be that  $p$ -scores translate to trait values in a non-linear way; or it might be that the relation between trait value and fitness is not linear. Here, we assume that, for either of these two reasons, fitnesses follow a quadratic model:

$$w_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 + \varepsilon_i \quad (\text{B})$$

where  $\alpha$  is the baseline fitness,  $\beta_1$  and  $\beta_2$  are the linear and the quadratic effect of the  $p$ -score on fitness, and  $\varepsilon_i$  is a noise term.<sup>2</sup>

Then we can, for example, take a parent population that consists of two individuals. The first has a  $p$ -score of 1, and two offspring; and the second has a  $p$ -score of  $\frac{1}{2}$ , and no offspring. In other words,  $p = \left[1, \frac{1}{2}\right]$ ,  $w = [2, 0]$  and

$$A = P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

The ingredients of the Price equation then are:  $\bar{w} = 1$ , as there are 2 parents and 2 offspring;  $\Delta \bar{p} = \frac{1}{4}$ , because the average  $p$ -score went from  $\frac{3}{4}$  to 1;  $E(w\Delta p) = 0$ , by definition with asexual reproduction without mutations;  $\text{Cov}(w, p) = \frac{1}{4}$ ,  $\text{Var}(p) = \frac{1}{16}$  and  $\beta = 4$ . The calculations are in Appendix B.

The Price equation therefore is:

$$\bar{w}\Delta \bar{p} = \beta \cdot \text{Var}(p) + E(w\Delta p)$$

$$1 \cdot \frac{1}{4} = 4 \cdot \frac{1}{16} + 0$$

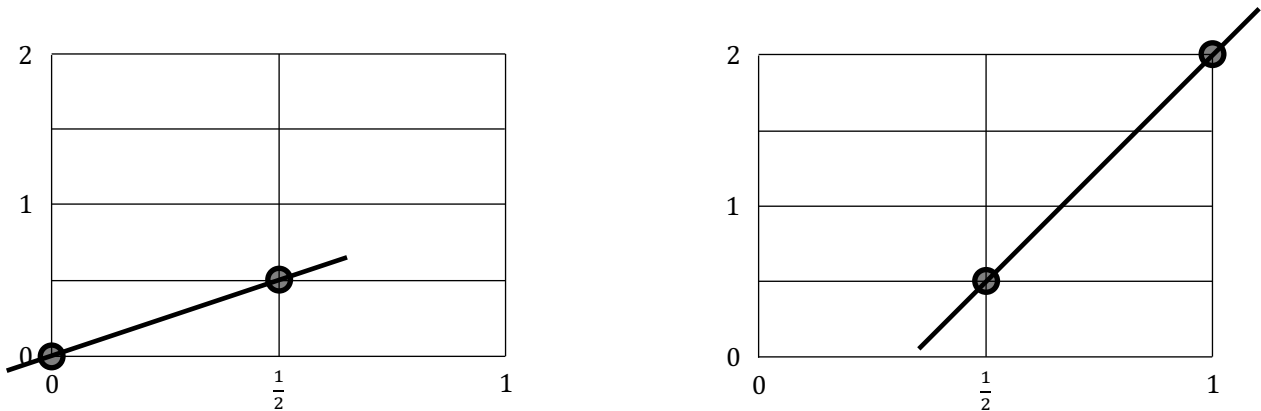
As was the case in the first example, the sample size here is too small to estimate anything with significance. In this case, there is moreover an additional problem. The  $\beta$  from the Price equation would be the right estimator to consider if the true model were to be model (A) from the previous example. If the true model is model (B), however, then with ever larger

---

<sup>2</sup> If we define  $w^0$ ,  $w^{\frac{1}{2}}$ , and  $w^1$  as the expected fitnesses of an individual with  $p$ -score 0,  $\frac{1}{2}$  and 1, then  $w^0 = \alpha$ ,  $w^{\frac{1}{2}} = \alpha + \frac{1}{2}\beta_1 + \frac{1}{4}\beta_2$ , and  $w^1 = \alpha + \beta_1 + \beta_2$ . This we can rewrite as  $\alpha = w^0$ ,  $\beta_1 = -3w^0 + 4w^{\frac{1}{2}} - w^1$  and  $\beta_2 = 2(w^0 - 2w^{\frac{1}{2}} + w^1)$ .

sample sizes, the data would get ever likelier to reveal that, which would imply that the  $\beta$  from the Price equation cannot be interpreted meaningfully for any population size.

A symptom of this is that even in a situation without noise, or with an infinitely large population, the  $\beta$  would vary, depending on the composition of the parent population. If  $\beta_1 = 0$  and  $\beta_2 = 2$ , and the parent population would only contain individuals with a  $p$ -score of 0 or a  $p$ -score of  $\frac{1}{2}$ , then the  $\beta$  in the Price equation would be 1 (see Fig. 1 on the left). This is true, regardless of the relative shares of individuals with a  $p$ -score of 0 and a  $p$ -score of  $\frac{1}{2}$ ; all that is required is that no individual with a  $p$ -score of 1 is present in the parent generation. If, on the other hand, the parent population would only contain individuals with a  $p$ -score of  $\frac{1}{2}$  or a  $p$ -score of 1, then the  $\beta$  in the Price equation would be 3 (see Fig. 1 on the right). If we apply what we find at the very end of Section 5 to this example, then we find that the  $\beta$  in the original Price equation equals  $\beta = 2 \cdot \frac{\text{Cov}(p, p^2)}{\text{Var}(p)}$ . This implies that the  $\beta$  depends on the composition of the population. If we furthermore assume random mating, this simplifies to  $\beta = 2 \cdot \left(\frac{1}{2} + \bar{p}\right)$ , where  $\bar{p}$  is the average  $p$ -score in the parent population. With  $\bar{p}$  ranging from 0 to 1, this implies that the  $\beta$  in the Price equation then ranges from 1 to 3.



**Fig. 1 | Dependency on the parent population.** We assume that there is no noise, or that the population is infinitely large, so that the average fitness matches the expected values. On the left a parent population, all members of which either have a  $p$ -score of 0, or a  $p$ -score of  $\frac{1}{2}$ . With expected fitnesses belonging to these  $p$ -scores of 0 and  $\frac{1}{2}$ , respectively, that results in a  $\beta$  of 1. On the right a parent population, all members of which either have a  $p$ -score of  $\frac{1}{2}$ , or a  $p$ -score of 1. With expected fitnesses belonging to these  $p$ -scores of  $\frac{1}{2}$  and 2, respectively, that results in a  $\beta$  of 3. In the literature, the dependence of the  $\beta$  on the parent population is sometimes referred to as *dynamical insufficiency*. That, however, is not what this is in this case, as here, this is really a symptom of misspecification. If the model that generated the data would have been linear (model A), then, absent the noise, the  $\beta$  would have been the same, regardless of the composition of the parent population. Dynamic (in)sufficiency is discussed in Section 6.

This implies that if the true underlying model is model (B), and not model (A), we cannot interpret  $\beta$  as the (linear) effect of these genes on fitness. There is however an alternative equation in which the regression coefficients would allow for a meaningful interpretation. In order to see what that would be, we go back to the derivation of the Price equation we went through earlier.

### Derivation of the generalized Price equation.

In the standard derivation of the Price equation, we arrived at

$$\bar{w}\Delta\bar{p} = \text{Cov}(w, p) + E(w\Delta p) \quad (\text{P.E.C})$$

where

$$\text{Cov}(w, p) = \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i$$

and

$$E(w\Delta p) = \frac{1}{n} \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right)$$

Now assume that we know that the fitnesses are generated by a model that can be written as follows

$$w_i = \alpha + \sum_{r=1}^R \beta_r p_i^r + \varepsilon_i$$

where  $\varepsilon_i$  is a noise term with expected value 0. If  $R = 1$ , this is the standard linear model we considered in Example 1.1;  $w_i = \alpha + \beta_1 p_i + \varepsilon_i$ . If  $R = 2$ , that is the quadratic model we are considering in Example 2.1;  $w_i = \alpha + \beta_1 \cdot p_i + \beta_2 \cdot p_i^2 + \varepsilon_i$ . Later, we will expand the set of models to all models that include a constant  $\alpha$  and a linear term  $\beta_1$ , but for the derivation it works perfectly fine to focus on this set of models.

If we have confidence in a model for some  $R \geq 1$ , we think that the true fitnesses are best described by

$$\hat{w}_i = \alpha + \sum_{r=1}^R \beta_r p_i^r = w_i - \varepsilon_i$$

for some choice of  $\alpha$  and the  $\beta_r$ 's. Therefore, it will be useful to think of this as the vector of estimated fitnesses, which we will indicate by

$$\hat{w} = [\hat{w}_1, \dots, \hat{w}_n]$$

Now for any  $R \geq 1$ , minimizing the sum of squared errors,

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \left( w_i - \left( \alpha + \sum_{r=1}^R \beta_r p_i^r \right) \right)^2$$

will imply that the derivatives to  $\alpha$  and to  $\beta_1$  are set to 0 (along with all other derivatives to the parameters of the model). This has two implications.

Setting the derivative to  $\alpha$  to 0 implies that we choose  $\alpha$  such that

$$-2 \sum_{i=1}^n \left( w_i - \left( \alpha + \sum_{r=1}^R \beta_r p_i^r \right) \right) = 0$$

and therefore also such that

$$\sum_{i=1}^n w_i = \sum_{i=1}^n \left( \alpha + \sum_{r=1}^R \beta_r p_i^r \right) = \sum_{i=1}^n \hat{w}_i$$

In other words, although for each individual  $i$  its actual number of offspring  $w_i$  and its estimated fitness  $\hat{w}_i$  may, and typically will, differ, if we arrive at  $\hat{w}_i$  by minimizing least squares with a model that includes  $\alpha$ , then they do add up to the same total.

Setting the derivative to  $\beta_1$  to 0 implies that we choose  $\beta_1$  such that

$$\sum_{i=1}^n -2p_i \left( w_i - \left( \alpha + \sum_{r=1}^R \beta_r p_i^r \right) \right) = 0$$

and therefore

$$\sum_{i=1}^n p_i w_i = \sum_{i=1}^n p_i \left( \alpha + \sum_{r=1}^R \beta_r p_i^r \right) = \sum_{i=1}^n p_i \hat{w}_i$$

In other words, for each individual the  $w_i$  and  $\hat{w}_i$  may differ (and if they do, then also  $p_i w_i$  and  $p_i \hat{w}_i$  will be different), but if we arrive at  $\hat{w}_i$  by minimizing least squares with a model that includes  $\beta_1$ , then the weighted sum  $\sum_{i=1}^n p_i w_i$  will nonetheless equal  $\sum_{i=1}^n p_i \hat{w}_i$ .

These two observations imply that if we use the data to estimate such a model with ordinary least squares (OLS), then we can also write

$$\text{Cov}(w, p) = \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i = \frac{1}{n} \sum_{i=1}^n \hat{w}_i p_i - \frac{1}{n^2} \sum_{i=1}^n \hat{w}_i \sum_{i=1}^n p_i = \text{Cov}(\hat{w}, p)$$

The Generalized Price equation in covariance form then becomes

$$\bar{w} \Delta \bar{p} = \text{Cov}(\hat{w}, p) + E(w \Delta p) \quad (\text{GPE.C})$$



Now we can focus on the term summarized as  $\text{Cov}(\hat{w}, p)$ , and fill in  $\hat{\alpha} + \sum_{r=1}^R \hat{\beta}_r p_i^r$  for  $\hat{w}_i$ . We write  $\hat{\alpha}$  and  $\hat{\beta}_r$  to indicate that these are not just any  $\alpha$  and  $\beta_r$ , but the ones we find by minimizing the sum of squared errors.

$$\frac{1}{n} \sum_{i=1}^n \left( \hat{\alpha} + \sum_{r=1}^R \hat{\beta}_r p_i^r \right) p_i - \frac{1}{n} \sum_{i=1}^n \left( \hat{\alpha} + \sum_{r=1}^R \hat{\beta}_r p_i^r \right) \frac{1}{n} \sum_{i=1}^n p_i$$

This can be shortened to

$$\frac{1}{n} \sum_{i=1}^n \left( \sum_{r=1}^R \hat{\beta}_r p_i^r \right) p_i - \frac{1}{n} \sum_{i=1}^n \left( \sum_{r=1}^R \hat{\beta}_r p_i^r \right) \frac{1}{n} \sum_{i=1}^n p_i$$

and if we change the summation order, this becomes

$$\sum_{r=1}^R \hat{\beta}_r \left( \frac{1}{n} \sum_{i=1}^n p_i^{r+1} - \frac{1}{n} \sum_{i=1}^n p_i^r \frac{1}{n} \sum_{i=1}^n p_i \right) = \sum_{r=1}^R \hat{\beta}_r \text{Cov}(p, p^r)$$

Therefore, if there is a model  $w_i = \alpha + \sum_{r=1}^R \beta_r p_i^r$  that we believe generated the data, and for which we estimate the parameters using ordinary least squares (OLS), we can always write the following equation, which we will call the Generalized Price equation in regression form. It is general, in the sense that it produces different equations for different choices of  $R$ , and therefore for every model within this set of models, where fitnesses are polynomials. In Section 3, we will see that by combining the Generalized Price equation in covariance form with an even richer set of models, we can generalize the Generalized Price equation in regression form even further.

$$\bar{w} \Delta \bar{p} = \sum_{r=1}^R \hat{\beta}_r \text{Cov}(p, p^r) + E(w \Delta p) \quad (\text{GPE.R1})$$

If  $R = 1$  we have the linear model,  $w_i = \alpha + \beta_1 p_i + \varepsilon_i$ , and since  $\text{Cov}(p, p) = \text{Var}(p)$  the above equation then becomes

$$\bar{w} \Delta \bar{p} = \hat{\beta}_1 \text{Var}(p) + E(w \Delta p) \quad (2.1)$$

If  $R = 2$  we have the quadratic model,  $w_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 + \varepsilon_i$ , and the above equation becomes

$$\bar{w} \Delta \bar{p} = \hat{\beta}_1 \text{Var}(p) + \hat{\beta}_2 \text{Cov}(p, p^2) + E(w \Delta p) \quad (2.2)$$

This moreover works for all models; the only thing that we need for this to work is that the statistical model includes a fixed term (for which we set the derivative of  $\alpha$  to 0) and a linear term (for which we set the derivative of  $\beta_1$  to 0). What is important to keep in mind, is that if we apply the linear model to a dataset, and then we apply the quadratic model to the same dataset, then the  $\hat{\beta}_1$ 's will be different between them. This is also what we encounter in our everyday statistics; if we estimate the same parameters in one model versus the other, the estimates will change.

Which of these equations has a meaningful interpretation, and/or to what degree, is perfectly in step with standard statistics. We may not have enough data to pick any model with any confidence, nor to estimate any parameter with significance. In that case none of these equations has a meaningful interpretation. We may, on the other hand, have an extremely large dataset that allows us to say with large confidence that the true model that generated them is for instance quadratic, and the sample size may also allow us to estimate the parameters with high accuracy. In that case equation (2.2) above has a meaningful interpretation, but not equation (2.1). Equations for  $R \geq 3$  will then typically only differ marginally from equation (2.2), with estimates for  $\hat{\beta}_i, i \geq 2$ , close to 0.

The standard Price equation in regression form is equation (2.1) above. This equation only has a meaningful interpretation if we have confidence that the data are generated by a linear model ( $R = 1$ ). Anything that one would be tempted to infer from the standard Price equation, without doing the statistics that confirms that the data are indeed generated by the linear model, is unwarranted. If a statistical test rejects model (A), then one cannot interpret the  $\beta$  in the standard Price equation as the effect of these genes on fitness.

Many researchers are enthusiastic about the fact that the Price equation is an identity. This regularly leads to the belief that the Price equation “cannot be wrong”. While it is correct to say that the Price equation is not wrong, in the sense that the left-hand side is equal to the right-hand side, it is important to also observe that this property is shared with other identities. In fact, we have found a range of different equations, one for every  $R \geq 1$ , and all of them are identities. If the “not being wrong” would just be about whether the left and the right-hand side are equal, then these equations can be (very) different, while none of them are wrong. If the “not being wrong” is to pertain to inferences one would draw from the Price equation, or interpretations of terms in it, then the multiplicity of equations implies that there *must* be some scope for being wrong, since these equations can differ from each other, leading to conclusions or interpretations that are at odds with each other.

The Price equation can be generalized, in the sense that one can write a variety of Price-like equations for a variety of possible true models, that may have generated the data. Which one we can interpret meaningfully with how much confidence depends on completely standard statistical considerations concerning model specification and significance of parameter estimates.

Before we go on to diploid species, there are a few more remarks that are worth making. The first is that I would like to reiterate that the genes that we are considering, and the genetic architecture, can limit which models make sense. If the  $p$ -score can only be 0 or 1, for instance, as it is in Section 2.1, then all models from the set we are considering here can be reduced to a model with  $R = 1$ . If we take

$$w_i = \alpha + \sum_{r=1}^R \beta_r p_i^r + \varepsilon_i$$

in combination with a binary  $p$ -score, then, since  $p_i^r = p_i$  for all  $r \geq 1$ , we can rewrite this as

$$w_i = \alpha + \left( \sum_{r=1}^R \beta_r \right) p_i + \varepsilon_i$$

This brings us back to a model with  $R = 1$ , if we choose  $\sum_{r=1}^R \beta_r$  as the coefficient for the linear term. Equivalently, if we look at the Generalized Price equation in regression form, then the fact that  $p^r = p$  also implies that  $\text{Cov}(p, p^r) = \text{Cov}(p, p) = \text{Var}(p)$  for all  $r$ , and hence the Generalized Price equation in regression form can be written as

$$\bar{w}\Delta\bar{p} = \sum_{r=1}^R \hat{\beta}_r \text{Cov}(p, p^r) + E(w\Delta p) = \left( \sum_{r=1}^R \hat{\beta}_r \right) \text{Var}(p) + E(w\Delta p)$$

This is the Generalized Price equation for the linear model, if we, again, choose  $\sum_{r=1}^R \beta_r$  as the coefficient for the linear term.

All of this is a formal way of saying that in this case, the regression coefficients only have meaning for the linear model, which is the model with  $R = 1$ .

The second remark is that we can also write the general regression form as follows.

$$\bar{w}\Delta\bar{p} = \left( \sum_{r=1}^R \hat{\beta}_r \frac{\text{Cov}(p, p^r)}{\text{Var}(p)} \right) \text{Var}(p) + E(w\Delta p) \quad (\text{GPE.R2})$$

The third remark is that for any model in the set of models we are considering here, if the error terms have a constant distribution with expectation 0 (that is, the distribution is the same for all values of the  $p$ -score), and the data are indeed generated by that model, then the Gauss-Markov theorem implies that the regression coefficients  $\hat{\beta}_r$  for that model are Best Linear Unbiased Estimators, which means that they have minimal variance within the set of unbiased estimators. It is however quite possible that the distribution of numbers of offspring depends on the  $p$ -score in ways that do not satisfy this.

One possibility is that the distribution of the error terms does not have a constant variance. An example would be if the number of offspring is drawn from a binomial distribution with  $N$  trials and success probability  $\frac{\alpha + \beta_1 p_i}{N}$ . This makes  $\alpha + \beta_1 p_i$  the expected number of offspring, and it makes  $w_i = \alpha + \beta_1 p_i + \varepsilon_i$  a well-specified model, but the variance of the error term now depends on  $p_i$ . The Gauss-Markov theorem then no longer applies, and the regression coefficients  $\hat{\beta}_r$  no longer have minimal variance. If the error terms still have expectation 0 – as they do in these examples – then they are however still unbiased. This implies that the setup of the Generalized Price equation is still useful. Applying OLS will still produce regression coefficients  $\hat{\beta}_r$  that are unbiased, even though there may be other unbiased estimation procedures that reduce the variance of the estimators. Heteroskedasticity (the variance of the error term not being constant) also makes a difference for what appropriate statistical tests are.

It is however also possible that the data are generated by a model that does not fit the set of models specified above. This would imply a departure from the setup that allows us to use

the Generalized Price equation, which is predicated on the statistical model including a constant term, and a term that is linear in the  $p$ -score.

### 3. The Price equation for diploid species.

We will now repeat this for a diploid, sexually reproducing species. This is also a special case of the original Price equation, which allows for any possible ploidy. What we add to the original Price equation, besides a bit of matrix notation, is that we separate both generations in females and males. This will be useful for the examples.

**Parents: mothers and fathers.** Let the parent generation consist of  $k$  females, numbered from  $i = 1$  to  $i = k$ , and  $n - k$  males, numbered from  $i = k + 1$  to  $i = n$ . This makes for a total of  $n$  individuals in the parent generation.

**Offspring: girls and boys.** Let the offspring generation consist of  $l$  females, numbered from  $j = 1$  to  $j = l$ , and  $m - l$  males, numbered from  $j = l + 1$  to  $j = m$ . This makes for a total of  $m$  individuals in the parent generation.

**Which kids belong to which parents.** Everyone in the offspring generation is the offspring of one female and one male parent in the parent generation. Who are the parents of which offspring, we represent, again, with an offspring matrix  $A$ ; let  $A_{ij} = 1$  if  $i$  is  $j$ 's parent, and  $A_{ij} = 0$  if not. Every offspring having one mother and one father means that  $\sum_{i=1}^k A_{ij} = 1$  for all  $j$  (one mother per child), and  $\sum_{i=k+1}^n A_{ij} = 1$  for all  $j$  (one father per child). This also implies  $\sum_{i=1}^n A_{ij} = 2$  for all  $j$  (two parents per child).

**Parent fitnesses.** The number of offspring of parent  $i$ , denoted by  $a_i$ , is found by summing over the offspring;  $a_i = \sum_{j=1}^m A_{ij}$ . The fitness of parent  $i$  we get by dividing this by the ploidy;  $w_i = \frac{1}{2} \sum_{j=1}^m A_{ij}$ . Because everyone in the offspring generation has two parents in the parent population, the sum of these fitnesses of the parents must equal the number of individuals in the offspring generation:  $\sum_{i=1}^n w_i = m$ .

**Genes in the parent generation.** The "dose" of a gene [1], or the  $p$ -score [7] per individual in the parent population, is denoted by  $p_i, i = 1, \dots, n$  for all parents. This dose may be restricted to be 0,  $\frac{1}{2}$ , or 1, if this is a counter for a specific allele, of which one can have 0, 1, or 2 copies. The idea of a dose or a  $p$ -score, however, is that this allows for any measure for the presence or absence of alleles all over the genome, and in particular genes that all contribute to a certain trait value. This means that  $p_i$  can take values on some subset of the interval  $[0,1]$ .

**Which genes are passed on.** The matrix  $P$  represents the  $p$ -score of the successful gametes that went into the offspring. This works as follows:  $P_{ij} = 0$  if individual  $i$  from the parent population is just not a parent of individual  $j$  in the offspring population ( $A_{ij} = 0$ ). If individual  $i$  from the parent population is a parent of individual  $j$  in the offspring population ( $A_{ij} = 1$ ), then  $P_{ij}$  is the  $p$ -score of the successful gamete. Because every individual is

composed of two gametes, this is a value in  $\left[0, \frac{1}{2}\right]$ . There may be restrictions on what  $P_{ij}$  can be depending on what  $p_i \in [0,1]$  is; for instance, if  $p_i$  counts alleles at one locus, as suggested above, then  $p_i$  can be either 0, if the gamete does not carry it, or  $\frac{1}{2}$  if it does. In this case, that would mean that if  $A_{ij} = 1$  ( $i$  is  $j$ 's parent), and the parent has 0 copies of the allele ( $p_i = 0$ ), then the gametes cannot contain any copies either ( $P_{ij} = 0$ ); if the parent has 1 copy of the allele ( $p_i = \frac{1}{2}$ ), then the gametes can either contain 0 or 1 copy ( $P_{ij} = 0$  or  $P_{ij} = \frac{1}{2}$ ); and if the parent has 2 copies of the allele ( $p_i = 1$ ), then the gametes must contain a copy as well ( $P_{ij} = \frac{1}{2}$ ). More generally, without mutations, the genetic details would impose restrictions that will typically include that  $p_i = 0$  will imply  $P_{ij} = 0$  for all of its offspring, and that  $p_i = 1$  implies  $P_{ij} = \frac{1}{2}$  for all of its offspring (that is, for all  $j$  for which  $A_{ij} = 1$ ).

**Genes in the offspring population.** The  $p$ -score of individual  $j$  in the offspring generation is the sum of the  $p$ -scores of its gametes;  $p'_j = \sum_{i=1}^n P_{ij}$ .

**Change.** The average  $p$ -score in the parent generation is  $\bar{p} = \frac{\sum_{i=1}^n p_i}{n}$ . The average  $p$ -score in the offspring generation is  $\bar{p}' = \frac{\sum_{j=1}^m p'_j}{m}$ .

**Derivation of the Price equation.** The key ingredient in the derivation of the Price equation is that one can calculate the sum of the  $p$ -scores in the offspring generation in two ways.

1. One can first calculate the  $p$ -score of every individual in the offspring generation by adding the  $p$ -scores of the incoming gametes of each offspring individual ( $p'_j = \sum_{i=1}^n P_{ij}$  for every offspring  $j$ ), and then sum these over all the members of the offspring generation. That amounts to  $\sum_{j=1}^m p'_j = \sum_{j=1}^m \sum_{i=1}^n P_{ij}$ .
2. One can also first add up the  $p$ -scores of the successful outgoing gametes of the parents ( $\sum_{j=1}^m P_{ij}$  for parent  $i$ ) and then sum over the parents. This amounts to a switch in summation order from  $\sum_{j=1}^m \sum_{i=1}^n P_{ij}$  to  $\sum_{i=1}^n \sum_{j=1}^m P_{ij}$ . This we can always do, because  $n$  and  $m$  are finite.

Despite the differences (combinations of  $A$  and  $P$  that fit the asexual model setup of Section 2 do not fit the sexual setup in this section, and vice versa), the remainder of the derivation is the exact same as the derivation for asexual reproduction in Section 2. Also the derivation of the generalized Price equation is the same.

In order to illustrate the Price equation for a diploid, sexually reproducing species, we have three more examples. These examples moreover further illustrate that the interpretation of the Price equation is bound by the exact same considerations that confine interpreting parameter estimates in normal statistics.

### Example 3.1: an allele with a fixed effect on fitness

We begin with the equivalent of Example 1.1. The only difference between the setting with asexual reproduction from Example 1.1 and the setting with sexual reproduction is that in the latter, there is scope for model misspecification. We nonetheless start by assuming that we know what the underlying model of fitness is, and that it is linear:

$$w_i = \alpha + \beta p_i + \varepsilon_i \quad (\text{A})$$

Here, the  $p$ -score  $p_i$  can be 0,  $\frac{1}{2}$ , or 1, depending on whether the individual has 0, 1, or 2 copies of a certain allele.

Then we take a parent population consisting of two mothers and two fathers. The first mother has two copies of the allele, the second has none, the first father has one copy of the allele, the second has none. In other words,  $p = [1, 0, \frac{1}{2}, 0]$ . Mother 1 has two kids with father 1 and one with father 2, mother 2 has one kid with father 2, which makes  $w = [\frac{3}{2}, \frac{1}{2}, 1, 1]$  and

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

The transmission matrix is

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This makes  $p' = [\frac{1}{2}, 1, \frac{1}{2}, 0]$ , which we get by summing over the columns. The ingredients of the Price equation then are:  $\bar{w} = 1$ , as there are 4 parents and 4 offspring;  $\Delta\bar{p} = \frac{1}{8}$ , because the average  $p$ -score went up from  $\frac{3}{8}$  to  $\frac{1}{2}$ ;  $E(w\Delta p) = 0$ , because the only parent with a non-binary  $p$ -score had a  $p$ -score of  $\frac{1}{2}$ , and this parent had one successful gamete without, and one with the allele;  $\text{Cov}(w, p) = \frac{1}{8}$ ; and  $\text{Var}(p) = \frac{11}{64}$ . In line with the insight from the Generalized Price equation, we give the  $\beta$  a hat, indicating that this is what the estimator would be for the true  $\beta$ , under the assumption that the true model is in fact linear. The  $\beta$  in the original Price equation is defined as  $\frac{\text{Cov}(w, p)}{\text{Var}(p)}$ , which is indeed the estimator of  $\beta$ , under the assumption of the data are generated by the linear model, so for the original Price equation, this does not change how the  $\beta$  is calculated. For this example, this is  $\hat{\beta} = \frac{8}{11}$ . The calculations are in Appendix C.

The Price equation, and the Generalized Price equation under the linear model, therefore is:

$$\bar{w}\Delta\bar{p} = \hat{\beta} \cdot \text{Var}(p) + E(w\Delta p)$$

$$1 \cdot \frac{1}{8} = \frac{8}{11} \cdot \frac{11}{64} - 0$$

As in previous examples, this sample size is not large enough to estimate the true  $\beta$  with any significance. Also, the sample size is not large enough to test if the model  $w_i = \alpha + \beta p_i + \varepsilon_i$  is accurately specified. In this case, the  $\hat{\beta}$  therefore cannot be interpreted as the fitness effect of the gene.

One can however also imagine a vector  $p$ , and matrices  $A$  and  $P$ , that together represent a much larger population, and therefore a much larger dataset. If this results in an estimate of the true  $\beta$  that is significantly different from 0; if a statistical test with sufficient power does not reject model (A); and if a statistical test with sufficient power rejects that the expectation of the noise term  $\varepsilon_i$  is not zero,<sup>3</sup> then the  $\beta$  in the Price equation does have a meaningful interpretation; it is an unbiased estimator of the true  $\beta$ , which is the effect of the  $p$ -score on fitness

The main observation here is the same as with Examples 2.1 and 2.2. The extent to which we can give a meaningful interpretation of the right-hand side of the Price equation is determined completely by standard statistical considerations. With small sample sizes,  $\hat{\beta}$  will typically not be significant, even if the true  $\beta$  would be different from 0. Also tests for linearity will have low power, and will not reject linearity, even if the true model would not be linear. Larger sample sizes allow us to infer more from the data, but also then, the interpretation is bound by the rules of statistics. Any interpretation will have to be done with the exact same reservations that one would have with standard statistics.

If the  $\hat{\beta}$  is significant, and statistical tests with sufficient power do not reject model (A), then the right-hand side of the original Price equation in regression form (which is the right-hand side of the Generalized Price equation in regression form for the model  $w_i = \alpha + \beta p_i + \varepsilon_i$ ) has a meaningful interpretation;  $\hat{\beta}$  is the estimated (linear) effect of the  $p$ -score on fitness.

### Example 3.2: heterozygote advantage / dominance

In this example, we begin with a description of a life cycle in which heterozygotes have an advantage. The quadratic model that describes this can, however, also be used to describe dominance for alternative choices of the parameters.

Suppose the  $p$ -score can be 0,  $\frac{1}{2}$ , or 1, and assume that for any of those  $p$ -scores there are equally many females as there are males. The life cycle is as follows. Those with a  $p$ -score of 0 all die. Those with a  $p$ -score of  $\frac{1}{2}$  all survive. Half of individuals with a  $p$ -score of 1 survive, the other half die. The female and male survivors then are randomly matched in pairs, and all of those pairs produce 3 kids.

---

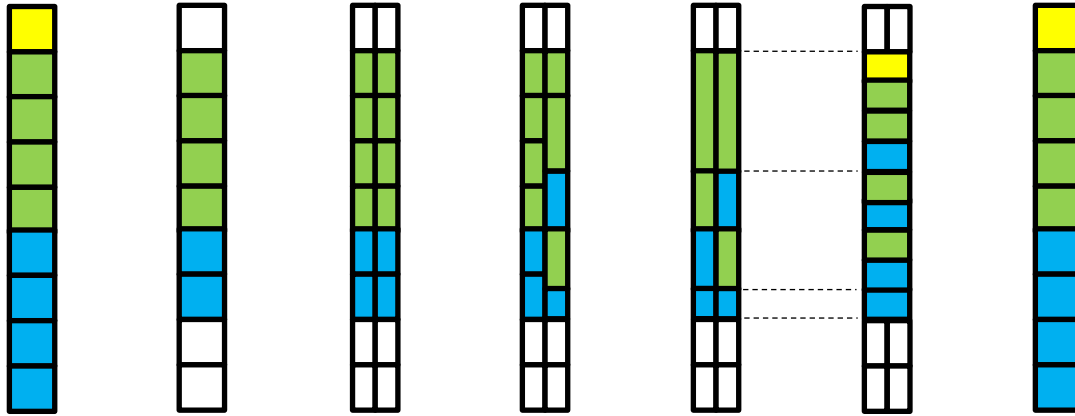
<sup>3</sup> This would depend on  $E(w\Delta p)$  being sufficiently close to 0 to reject that there is no fair meiosis.

The fitnesses for these three  $p$ -scores therefore are  $0$ ,  $\frac{3}{2}$ , and  $\frac{3}{4}$ . Another way to describe that is:

$$w_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 + \varepsilon_i \quad (\text{B})$$

where  $\varepsilon_i$  is a noise term, and  $\alpha = 0$ ,  $\beta_1 = \frac{21}{4}$ , and  $\beta_2 = -\frac{9}{2}$ .<sup>4</sup>

Then we can assume an infinite population, and we can assume that the shares in the parent population are as follows: 1 out of 9 parents has a  $p$ -score of 0; 4 out of 9 parents have a  $p$ -score of  $\frac{1}{2}$ ; and 4 out of 9 parents have a  $p$ -score of 1. These frequencies are chosen to be stable; the life cycle as described above creates an offspring generation that is identical to the parent population. Figure 2 below illustrates all steps in the calculations, which are in Appendix D, and that confirm that this is indeed a fixed point of the dynamics.



**Fig. 2 | A graphical representation of the life cycle for the equilibrium state of the population.** *Step 1: differential survival.* All individuals in the parent generation with a  $p$ -score of 0 die, and half of those with a  $p$ -score of 1 do. None of the parents with a  $p$ -score of  $\frac{1}{2}$  die, so the heterozygote has the highest fitness. *Step 2: mothers and fathers.* Half of all surviving parents are female, and half are male. *Step 3 and 4: random matching.* The parents are randomly matched. *Step 5: fair meiosis.* In each pair, the offspring inherits one allele from either parent. *Step 6: scaling up.* Parent pairs have on average 3 kids.

One can do this for non-equilibrium parent populations too, in which case the numbers will change, but it might be instructive to first focus on this equilibrium population state. For this state, the ingredients of the Price equation are:  $\bar{w} = 1$ , as the number of offspring equals the number of parents;  $\Delta \bar{p} = 0$ , because the frequencies do not change;  $E(w\Delta p) = 0$ ,

<sup>4</sup> In general,  $w^0 = \alpha$ ,  $w^{\frac{1}{2}} = \alpha + \frac{1}{2}\beta_1 + \frac{1}{4}\beta_2$ , and  $w^1 = \alpha + \beta_1 + \beta_2$  (see also footnote 1). If fitness is linear, that would mean  $\beta_2 = 0$ . Heterozygote advantage would be  $-2\beta_1 < \beta_2 < -\frac{2}{3}\beta_1$  (which can only hold for  $\beta_1 > 0$ ). A dominant gene would make  $w^{\frac{1}{2}}$  and  $w^1$  equal, and therefore  $\beta_2 = -\frac{2}{3}\beta_1$ ; a recessive gene would make  $w^0$  and  $w^{\frac{1}{2}}$  equal, and therefore  $\beta_2 = -2\beta_1$ .



because of fair meiosis in an infinite population;  $\text{Cov}(w, p) = 0$ ,  $\text{Var}(p) = \frac{1}{9}$  and  $\hat{\beta} = 0$ . The calculations are in Appendix D.

The Price equation, and the Generalized Price equation under the linear model, for the equilibrium population state therefore is:

$$\begin{aligned}\bar{w}\Delta\bar{p} &= \hat{\beta} \cdot \text{Var}(w, p) + E(w\Delta p) \\ 1 \cdot 0 &= 0 \cdot \frac{1}{9} + 0\end{aligned}$$

If we were to interpret the  $\hat{\beta}$  as the effect of the gene on fitness, then the fact that  $\beta = 0$  would suggest that the gene has no effect on fitness. This would be wrong, because which  $p$ -score an individual has clearly matters for fitness; it is just that the effect is not linear. At this particular frequency, it happens to be the case that the advantageous effect of an additional allele compared to the homozygote with a  $p$ -score of 0, and the disadvantageous effect of an additional allele compared to the homozygote with a  $p$ -score of  $\frac{1}{2}$  balance out. If the data are generated by Model (B), as we assume they are in this example, then one cannot interpret the  $\hat{\beta}$  from the Price equation as the effect of the gene on fitness, which one can do if the data were generated by Model (A). Moreover, as in example 2.2, the  $\hat{\beta}$  from the Price equation will depend on the composition of the parent population; for all parent populations that are not the equilibrium shares given above, the  $\hat{\beta}$  will be non-zero, and can be positive or negative. This is a symptom of misspecification. (It can also be a reflection of frequency dependence, as we will see in the next example, but here fitnesses are not frequency dependent).

With an infinitely large population, the Price equation would have a meaningful interpretation, if the true model were model (A). If the true model is model (B), however, even if the parent population is assumed to be infinitely large (which removes statistical considerations concerning sample size), the  $\hat{\beta}$  in the Price equation does not have a meaningful interpretation, due to misspecification of the model.

If the fitnesses depend on the  $p$ -scores in a quadratic way, as they do in this example, then, provided that we have sufficiently much data, a standard statistical exercise is likely to uncover the true model, and estimate its parameters with some accuracy and confidence. In this model, the fitnesses do not depend on the composition of the parent population, which means that for instance the expected values of the estimators for  $\beta_1$  and  $\beta_2$  do not depend on the composition of the parent population. The  $\hat{\beta}$  that we are likely to find in the original Price equation in regression form, which is the Generalized Price equation for the model  $w_i = \alpha + \beta p_i + \varepsilon_i$ , on the other hand, does depend on the composition of the parent population. This is caused by misspecification, because in this example we assume that the true model is  $w_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 + \varepsilon_i$ .

### Example 3.3: sex-determining genes

In this example, the  $p$ -score will represent sex determining genes. If we think of females as  $XX$  and males as  $XY$ , and we count the number of  $X$ 'es, then the  $p$ -score for females is 1 and for males it is  $\frac{1}{2}$  (everything would also work if the  $p$ -score would count the number of  $Y$ 's, in which case the  $p$ -score would be 0 for females and  $\frac{1}{2}$  for males). We can then write the Price equation, just based on the numbers of females and males in both generations.

The ingredients of the Price equation then are:  $\bar{w} = \frac{m}{n}$ , which is the size of the offspring generation over the size of the parent population;  $\Delta\bar{p} = \frac{1}{2}\left(\frac{l}{m} - \frac{k}{n}\right)$ , which is one half times the difference between the frequency of females in the offspring generation and the parent generation;  $E(w\Delta p) = \frac{1}{2n}\left(l - \frac{1}{2}m\right)$ , which does not depend on the sex ratio in the parent population;  $\text{Cov}(w, p) = \frac{m}{2n}\left(\frac{1}{2} - \frac{k}{n}\right)$ ,  $\text{Var}(p) = \frac{k}{4n}\left(1 - \frac{k}{n}\right)$  and  $\hat{\beta} = m \cdot \frac{n-2k}{k(n-k)}$ . The calculations are in Appendix E.

The Price equation therefore is:

$$\bar{w}\Delta\bar{p} = \hat{\beta} \cdot \text{Var}(p) + E(w\Delta p)$$

$$\frac{m}{n} \cdot \frac{1}{2} \left( \frac{l}{m} - \frac{k}{n} \right) = m \cdot \frac{n-2k}{k(n-k)} \cdot \frac{1}{4} \frac{k}{n} \left( 1 - \frac{k}{n} \right) + \frac{1}{2n} \left( l - \frac{1}{2}m \right)$$

The most interesting observation here is that the  $\hat{\beta}$  is determined entirely by the composition of the parent population (the  $k$  and the  $n$ ) and the number of offspring  $m$ . Just by the mechanics of sexual reproduction, if the number of females is smaller than the number of males ( $k < n - k \rightarrow 2k < n$ ), the  $\hat{\beta}$  will be positive, whereas it will be negative if the number of females is larger than the number of males.

In this example, the fitnesses are frequency dependent. Recovering this in a statistical exercise requires observations for different compositions of the parent population, and writing the Price equation for one composition of the parent population only is not informative.

These examples for a diploid, sexually reproducing species explore a bit more what the scope is for possible ways in which fitnesses can depend on  $p$ -scores. The message however remains the same. The standard Price equation is just one out of a range of Price-like equations. There is one for every model, and the standard Price equation is the one for the linear model. Which of those equations has a meaningful interpretation, and to what extent, depends completely on what statistics has to say about the possible underlying true models that could have generated the data.

#### 4. The Price equation with a $p$ -score and a $q$ -score.

Here we take the same setup as in Section 2, but now we add information about genes that are not included in the  $p$ -score, while they may have an effect on fitness. Besides a vector  $p$  of  $p$ -scores for the parents, a vector  $p'$  of  $p$ -scores for the offspring, and a transmission matrix  $P$ , we now also have a vector  $q$  of  $q$ -scores for the parents, a vector  $q'$  of  $q$ -scores for the offspring, and a transmission matrix  $Q$ , all of which satisfy the same properties that a genetic system of transmission imposes (such as, for example, that if  $A_{ij} = 0$ , then  $Q_{ij} = 0$ ).

In this setting, the Price equation tracks changes in the average  $p$ -score, while these  $p$ -scores are not the only genetic determinant of fitness. This enriches the set of possible true models, and therefore the set of possible Price equations, and it allows us to give additional illustrations of how there is no way around classical statistics if our aim is to find the right Price equation in this multitude of options.

#### Derivation of the Generalized Price equation in regression form for this richer set of models.

With this richer setup, allowing for  $p$ -scores and  $q$ -scores, we can choose a richer set of models that we consider. Assume, therefore, that we know that the fitnesses are generated by some model that can be written as follows:

$$w_i = \sum_{(k,l) \in E} \beta_{k,l} p_i^k q_i^l + \varepsilon_i$$

where  $\varepsilon_i$  is a noise term with expected value 0, and  $E$  is some set that indicates which coefficients  $\beta_{k,l}$  are nonzero. We assume that  $(0,0)$  and  $(1,0)$  are included. If those are the only ones in  $E$ , then this is the standard linear model we considered in Example 2.1;  $w_i = \alpha + \beta p_i + \varepsilon_i$ , with  $\beta_{0,0} = \alpha$  and  $\beta_{1,0} = \beta$ . If  $E = \{(0,0), (1,0), (2,0)\}$ , that is the quadratic model we are considering in Example 2.2;  $w_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 + \varepsilon_i$ , with  $\beta_{0,0} = \alpha$ ,  $\beta_{1,0} = \beta_1$ , and  $\beta_{2,0} = \beta_2$ . If  $E = \{(0,0), (1,0), (0,1), (1,1)\}$ , that is a model in which both the  $p$ -score and the  $q$ -score matter, and in which there is an interaction term;  $w_i = \beta_{0,0} + \beta_{1,0} p_i + \beta_{0,1} q_i + \beta_{1,1} p_i q_i + \varepsilon_i$ .

If we repeat what we did in Section 2, choosing parameter values that minimize the sum of squared errors, where  $(0,0)$  and  $(1,0)$  are included, we arrive at the same generalized Price equation in covariance form.

$$\bar{w} \Delta \bar{p} = \text{Cov}(\hat{w}, p) + E(w \Delta p) \quad (\text{GPE.C})$$

Also the continuation from the covariance form to the regression form is almost the same. For this, we focus on the term summarized as  $\text{Cov}(\hat{w}, p)$ .

$$\text{Cov}(\hat{w}, p) = \frac{1}{n} \sum_{i=1}^n \hat{w}_i p_i - \frac{1}{n^2} \sum_{i=1}^n \hat{w}_i \sum_{i=1}^n p_i$$

Here we fill in  $\sum_{(k,l) \in E} \hat{\beta}_{k,l} p_i^k q_i^l$  for  $\hat{w}_i$

$$\frac{1}{n} \sum_{i=1}^n \left( \sum_{(k,l) \in E} \hat{\beta}_{k,l} p_i^k q_i^l \right) p_i - \frac{1}{n} \sum_{i=1}^n \left( \sum_{(k,l) \in E} \hat{\beta}_{k,l} p_i^k q_i^l \right) \frac{1}{n} \sum_{i=1}^n p_i$$

Because  $p_i^0 q_i^0 = 1$ , which makes  $\frac{1}{n} \sum_{i=1}^n p_i^0 q_i^0 p_i = \frac{1}{n} \sum_{i=1}^n p_i = p_i^0 q_i^0 \frac{1}{n} \sum_{i=1}^n p_i$ , we can leave out the first term, which means this is also

$$\frac{1}{n} \sum_{i=1}^n \left( \sum_{(k,l) \in E \setminus (0,0)} \hat{\beta}_{k,l} p_i^k q_i^l \right) p_i - \frac{1}{n} \sum_{i=1}^n \left( \sum_{(k,l) \in E \setminus (0,0)} \hat{\beta}_{k,l} p_i^k q_i^l \right) \frac{1}{n} \sum_{i=1}^n p_i$$

and if we change the summation order, this becomes

$$\sum_{(k,l) \in E \setminus (0,0)} \hat{\beta}_{k,l} \left( \frac{1}{n} \sum_{i=1}^n p_i^{k+1} q_i^l - \frac{1}{n} \sum_{i=1}^n p_i^k q_i^l \frac{1}{n} \sum_{i=1}^n p_i \right) = \sum_{(k,l) \in E \setminus (0,0)} \hat{\beta}_{k,l} \text{Cov}(p, p^k q^l)$$

Therefore, if there is any model  $w_i = \sum_{(k,l) \in E} \beta_{k,l} p_i^k q_i^l$  that we believe generated the data, and for which we estimate the parameters using ordinary least squares (OLS), we can always write the following equation, which we call the Generalized Price equation in regression form, for this larger set of models.

$$\bar{w} \Delta \bar{p} = \sum_{(k,l) \in E} \hat{\beta}_{k,l} \text{Cov}(p, p^k q^l) + E(w \Delta p) \quad (\text{GPE.R3})$$

Because  $\text{Cov}(p, p^0 q^0) = \text{Cov}(p, 1) = 0$ , it, again, does not matter if we leave  $(0,0)$  in or take it out.

This first of all encompasses the models that were included in Section 2.

Choosing  $E = \{(0,0), (1,0)\}$  here is the same as choosing  $R = 1$  in Section 2; both give us the linear model,  $w_i = \alpha + \beta_{1,0} p_i + \varepsilon_i$ , and since  $\text{Cov}(p, p) = \text{Var}(p)$  the above equation becomes

$$\bar{w} \Delta \bar{p} = \hat{\beta}_{1,0} \text{Var}(p) + E(w \Delta p) \quad (4.1)$$

Choosing  $E = \{(0,0), (1,0), (2,0)\}$  here is the same as choosing  $R = 2$  in Section 2; both give us the quadratic model,  $w_i = \beta_{0,0} + \beta_{1,0} p_i + \beta_{2,0} p_i^2 + \varepsilon_i$ , and the above equation becomes

$$\bar{w} \Delta \bar{p} = \hat{\beta}_{1,0} \text{Var}(p) + \hat{\beta}_{2,0} \text{Cov}(p, p^2) + E(w \Delta p) \quad (4.2)$$

This setup also allows for models that are not included in Section 2.

Choosing  $E = \{(0,0), (1,0), (0,1), (1,1)\}$ , for instance, give us a model with an interaction term between the  $p$ -score and the  $q$ -score,  $w_i = \beta_{0,0} + \beta_{1,0} p_i + \beta_{0,1} q_i + \beta_{1,1} p_i q_i + \varepsilon_i$ , and the above equation becomes

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p) + \hat{\beta}_{0,1}\text{Cov}(p, q) + \hat{\beta}_{1,1}\text{Cov}(p, pq) + E(w\Delta p) \quad (4.3)$$

This works for all models; the only thing that we need for this to work is that the statistical model includes a fixed term (for which we set the derivative to its coefficient to 0) and a linear term (for which we set the derivative to its coefficient to 0).

Also here, we can rewrite the regression form:

$$\bar{w}\Delta\bar{p} = \left( \sum_{(k,l) \in E} \hat{\beta}_{k,l} \frac{\text{Cov}(p, p^k q^l)}{\text{Var}(p)} \right) \text{Var}(p) + E(w\Delta p) \quad (\text{GPE.R4})$$

The Price equation can be generalized, also in a setting with two sets of genes, one captured by the  $p$ -score, the changes in which we track, and one by a  $q$ -score, the changes in which we do not track. The larger variety of possible true models, compared to models with a  $p$ -score only, is reflected in the larger variety of Price-like equations. Which one we can interpret meaningfully with how much confidence depends, as before, on completely standard statistical considerations concerning model specification and significance of parameter estimates.

## 5. Modeling

Part of statistics is having to choose between statistical models, based on data. That means that the best we can do is to maximize the probability that we pick the right model. Modelling, on the other hand, relieves us of that uncertainty, as the model is just whatever we choose it to be. That makes it even more clear that there is something to be gained from matching the model that we happen to have chosen with the right Price-like equation.

While we have seen that the literature is somewhat sloppy on how to use the Price equation on data, it is also not always very specific or precise on how the Price equation is to be used on models. In this section, we will first circumvent all complications, by assuming that we have a deterministic model, without noise. This is of course totally unrealistic, but it allows us to make a point about over- and underspecification of Price-like equations in a modelling context. This point is similar to the point about over- or underspecification in a statistical context. Moreover, this point also carries over to a setting where we assume an infinitely large population. This assumption allows us to treat the population dynamics as deterministic, even if there is uncertainty at the individual level about the number of offspring. In the last subsection, we will drop those assumptions, and show how even well-specified Price-like equations have limitations in describing properties of dynamics in stochastic models that do not assume that the population size is large enough for models with infinite population size to be a good approximation.

For simplicity, we begin assuming asexual reproduction. All points made below carry over to a setting with sexual reproduction but are easier made without. At the end of this section, we will revisit heterozygote advantage, for which we do switch to sexual reproduction.

## Model A

The first model is  $w_i = \alpha + \beta_{1,0}p_i$ . The absence of a noise term severely limits what values  $\alpha$  and  $\beta_{1,0}$  can take. If  $p_i$  is binary, then the fact that numbers of offspring are integers means that we can also only choose integers for  $\alpha$  and  $\beta_{1,0}$ . If  $p_i$  is restricted to be  $0, \frac{1}{2}$  or  $1$ , then  $\alpha$  can still only be an integer, while  $\beta_{1,0}$  needs to be an even number. Larger sets of possible  $p$ -scores further reduce what values for  $\beta_{1,0}$  we can choose. If we assume infinite populations, with uncertainty at the individual level, but deterministic population dynamics, these restrictions are lifted, but for now we just assume that they are satisfied.

A deterministic model implies that for any given parent population, represented by a vector of  $p$ -scores  $p$ , there is only one possible transition to an offspring population. For that one transition, we can use the Generalized Price equation, where the sum of squared errors is minimized, and equal to 0, at  $\hat{\alpha} = \alpha$  and  $\hat{\beta}_{1,0} = \beta_{1,0}$ . Here we simply use the matching statistical model, which in this case only needs to include  $\alpha$  and  $\beta_{1,0}$ . The matching Price-like equation for this model is

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p)$$

where  $\hat{\beta}_{1,0}$  does not depend on which  $p$  we choose, and equals  $\beta_{1,0}$  for any composition of the parent population. This equation can therefore serve as a useful summary of the dynamics induced by the model.

## Model B

The second model is  $w_i = \alpha + \beta_{1,0}p_i + \beta_{0,1}q_i$ . Again, the deterministic nature of the model puts restrictions on the values we can choose for  $\alpha$ ,  $\beta_{1,0}$ , and  $\beta_{0,1}$ , but those restrictions would be lifted if we assume infinite populations, with uncertainty at the individual level, but deterministic dynamics at the population level.

As before, we can use the Generalized Price equation, where the sum of squared errors is minimized, and equal to 0, at  $\hat{\alpha} = \alpha$ ,  $\hat{\beta}_{1,0} = \beta_{1,0}$ , and  $\hat{\beta}_{0,1} = \beta_{0,1}$ . The matching Price-like equation for this model is

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p) + \hat{\beta}_{0,1}\text{Cov}(p, q) \quad (5.B)$$

in which  $\hat{\beta}_{1,0}$  and  $\hat{\beta}_{0,1}$  do not depend on the choice of  $p$  and  $q$ , and are equal to  $\beta_{1,0}$  and  $\beta_{0,1}$ , respectively, for any composition of the parent population.

## Model C

The third model is  $w_i = \alpha + \beta_{1,0}p_i + \beta_{0,1}q_i + \beta_{1,1}p_iq_i$ . Here we jump over the details, and go straight to the matching Price-like equation for this model, which is

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p) + \hat{\beta}_{0,1}\text{Cov}(p, q) + \hat{\beta}_{1,1}\text{Cov}(p, pq) \quad (5.C)$$

These three examples together illustrate that, also when modeling, there is a multitude of Price-like equations, that capture properties of the population dynamics implied by the model.

### Mismatching models and Price-like equations

It is important to note that also in this modelling setting, all of these Price-like equations remain identities, and they can be combined, not just with the model with respect to which they minimize the sum of squared errors, but also with all other, possibly richer models, as long as they have a constant term, and a linear term for  $p_i$ . As a symptom of such a mismatch, the coefficients in the Price-like equation stop being constants; they will vary with the composition of the parent population (the  $p$  and, if present, the  $q$  vector) if they are applied to other models. Another symptom is that the coefficients in the Price-like equation stop matching the according coefficients in the model.

#### *Applying the Generalized Price equation for Model A to Model B*

If we apply the Generalized Price equation for Model A to Model B, then the sum of squared errors can no longer be reduced to be 0, and is minimized at  $\hat{\beta}_{1,0} = \beta_{1,0} + \beta_{0,1} \frac{\text{Cov}(p,q)}{\text{Var}(p)}$ . The fact that the sum of squared errors is positive is noteworthy. The model is deterministic, and has no errors, and yet the sum of squared errors in the Price-like equation is larger than 0. By lack of actual errors, this can only be caused by misspecification of the model. The value for  $\hat{\beta}_{1,0}$  at which the sum of squared errors is minimized if we apply the Generalized Price equation for Model A to Model B follows from the fact that both  $\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p)$  and  $\bar{w}\Delta\bar{p} = \beta_{1,0}\text{Var}(p) + \beta_{0,1}\text{Cov}(p, q)$  are identities. The first follows from the derivation of the Generalized Price equation, combined with Model A, and the second was established above for transitions generated by Model B. We therefore end up with the following identity:

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p)$$

where  $\hat{\beta}_{1,0} = \beta_{1,0} + \beta_{0,1} \frac{\text{Cov}(p,q)}{\text{Var}(p)}$ . This implies that  $\hat{\beta}_{1,0}$  no longer is a constant, as it varies with  $p$  and  $q$  through the  $\frac{\text{Cov}(p,q)}{\text{Var}(p)}$  - term. It also implies that  $\hat{\beta}_{1,0}$  is not the linear effect for this model, which would be  $\beta_{1,0}$ .

#### *Applying the Generalized Price equation for Model A to Model C*

If we apply the Generalized Price equation for Model A to Model C, then the sum of squared errors can also no longer be reduced to be 0, and is minimized at  $\beta_{1,0} + \beta_{0,1} \frac{\text{Cov}(p,q)}{\text{Var}(p)} + \beta_{1,1} \frac{\text{Cov}(p,pq)}{\text{Var}(p)}$ . We therefore end up with the following identity:

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p)$$

where  $\hat{\beta}_{1,0} = \beta_{1,0} + \beta_{0,1} \frac{\text{Cov}(p,q)}{\text{Var}(p)} + \beta_{1,1} \frac{\text{Cov}(p,pq)}{\text{Var}(p)}$ . This implies that  $\hat{\beta}_{1,0}$  is no longer a constant, as it varies with  $p$  and  $q$  through the  $\frac{\text{Cov}(p,q)}{\text{Var}(p)}$  – term and the  $\frac{\text{Cov}(p,pq)}{\text{Var}(p)}$  – term. It also implies that  $\hat{\beta}_{1,0}$  is not the linear effect for this model, which would be  $\beta_{1,0}$ .

### *Applying the Generalized Price equation for Model B to Model C*

This would lead to more complex relations between, on the one hand,  $\hat{\beta}_{1,0}$  and  $\hat{\beta}_{0,1}$ , and, on the other, the true  $\beta_{1,0}$ ,  $\beta_{0,1}$ , and  $\beta_{1,1}$ , but the idea is the same: misspecification leads to a positive sum of squared errors, when it is being minimized, and coefficients that change with the state of the parent population.

Also with modeling, misspecification is possible. This leads to Price-like equations in which the coefficients are not constant.

### **Expected changes**

If we do away with the assumption of a deterministic model, one could still hope for the (Generalized) Price equation to be correct in expectation. In order to explore that possibility, we go back to Model A, but now with noise. That means that we go back to writing the model as  $w_i = \alpha + \beta_{1,0}p_i + \epsilon_i$ . Explicitly or implicitly, the noise is typically assumed to have mean zero, which makes  $\alpha + \beta_{1,0}p_i$  the expected number of offspring of someone with  $p$ -score of  $p_i$ . Numbers of realized offspring are obviously integers, limiting what values  $w_i$  can take, but the fact that we assume (continuous) probabilities with which an individual has these integer numbers of offspring, implies that  $\alpha$  and  $\beta_{1,0}$  can be any number. The model then captures how the  $p$ -score affects the distribution of the number of offspring, summarized by  $\alpha + \beta p_i$  being the expected value of that distribution.

We begin by fixing a parent generation characterized by a vector of  $p$ -scores;  $p = [p_1, \dots, p_n]$ . Any choice for  $p$  also implies an average  $p$ -score  $\bar{p}$  and a “variance” in  $p$ -scores  $\text{Var}(p)$ . The new generation then is a combination of  $n$  independent random variables, one for every individual, as given by the model. For any realization of these random variables, one could write the Generalized Price equation that belongs to Model A. With asexual reproduction and no mutation – which implies that  $E(w\Delta p) = 0$  for every realization of these random variables – this would be

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0}\text{Var}(p)$$

It is worth stressing once more that this equation holds for every realization of the random variables. The average fitness  $\bar{w}$  and the value of  $\hat{\beta}_1$  do, of course, depend on the realization, while the expected values of  $\bar{w}$  and  $\hat{\beta}_{1,0}$ , obviously, do not. The expected value of  $\bar{w}$  is the expected average fitness  $\frac{1}{n} \sum_{i=1}^n (\alpha + \beta_{1,0}p_i)$ . The expected value of  $\hat{\beta}_{1,0}$  is  $\beta_{1,0}$ , because we have assumed that the data are generated by Model A, and that is what it means for  $\hat{\beta}_{1,0}$  to be an unbiased estimator of  $\beta_{1,0}$ . All of this could make one hope that maybe the



(Generalized) Price equation could reflect a property of the *expected* change in average  $p$ -score implied by the model. This hope is unfounded, as the following simple example shows.

Assume a population of two parents, one without and one with a certain gene;  $p = [0,1]$ . Without the gene, the probability of having 0 offspring and the probability of having 2 is both  $\frac{1}{2}$ . With the gene, the probability of having 1 offspring and the probability of having 3 is both  $\frac{1}{2}$ . That implies that the model fits  $w_i = \alpha + \beta_{1,0}p_i + \epsilon_i$ , where  $\alpha = 1$ ,  $\beta_{1,0} = 1$ , and  $\epsilon_i$  is 1 or  $-1$ , both with probability  $\frac{1}{2}$ .

This super simple model implies that there are four possible transitions, each with probability  $\frac{1}{4}$ .  $\text{Var}(p)$  is a property of the parent population, so  $\text{Var}(p) = \frac{1}{4}$  for all four transitions. Below we list the ingredients of the Price equation in regression form for the 4 transitions, all of which happen with probability  $\frac{1}{4}$ .

If the parents have 0 and 1 offspring, respectively, then  $\bar{w} = \frac{1}{2}$ ,  $\Delta\bar{p} = \frac{1}{2}$ , and  $\hat{\beta}_{1,0} = 1$ .

If the parents have 0 and 3 offspring, respectively, then  $\bar{w} = \frac{3}{2}$ ,  $\Delta\bar{p} = \frac{1}{2}$ , and  $\hat{\beta}_{1,0} = 3$ .

If the parents have 2 and 1 offspring, respectively, then  $\bar{w} = \frac{3}{2}$ ,  $\Delta\bar{p} = -\frac{1}{6}$ , and  $\hat{\beta}_{1,0} = -1$ .

If the parents have 2 and 3 offspring, respectively, then  $\bar{w} = \frac{5}{2}$ ,  $\Delta\bar{p} = \frac{1}{10}$ , and  $\hat{\beta}_{1,0} = 1$ .

This helps verify that the expected value of  $\hat{\beta}_{1,0}$  is indeed  $\beta_{1,0} = 1$ , as  $E[\hat{\beta}_{1,0}] = \frac{1}{4}(1 + 3 - 1 + 1) = 1$ . Also, that the expected value of  $\bar{w}$  is indeed  $\frac{1}{n}\sum_{i=1}^n(\alpha + \beta_{1,0}p_i) = \frac{1}{2}(1 + 2) = \frac{3}{2}$ , as  $E[\bar{w}] = \frac{1}{4}\left(\frac{1}{2} + \frac{3}{2} + \frac{3}{2} + \frac{5}{2}\right) = \frac{3}{2}$ . The expected value of  $\Delta\bar{p}$ , on the other hand, is  $E[\Delta\bar{p}] = \frac{1}{4}\left(\frac{1}{2} + \frac{1}{2} - \frac{1}{6} + \frac{1}{10}\right) = \frac{7}{30}$ . That implies that, while the (Generalized) Price equation holds for every realization, we cannot replace all terms in it by their expected values, as

$$\frac{3}{2} \cdot \frac{7}{30} \neq 1 \cdot \frac{1}{4}$$

In other words, it is not generally true that

$$E[\bar{w}]E[\Delta\bar{p}] = E[\hat{\beta}_{1,0}]\text{Var}(p)$$

with expectations taken over the random variables regarding the reproduction.

As a side note, we can try to reconcile this with [8], which contains a claim that one might think, at first sight, contradicts this. Therefore, we would have to observe that in that paper, there are two ways in which the Price equation is used differently. The first is that it uses the covariance form, and not the regression form. The second is that it does not use realized fitnesses, but defines a new variable  $v_i = \frac{w_i}{\bar{w}}$  for every realization, and uses that instead of the realized fitnesses. Both of these changes are needed to make this work; below we will show that neither one of the two is enough on its own. It is also very important to observe

that the change in variable from  $w_i$  to  $v_i$  obstructs any meaningful interpretation of the regression coefficients that is stable across realizations. The claim in [8] is therefore both true and useless.

For all 4 realizations, we listed  $\bar{w}$ ,  $\Delta\bar{p}$ , and  $\hat{\beta}_{1,0}$ , as calculated above, and we calculate  $\text{Cov}(p, w)$ . The regression coefficient for  $v$  instead of  $w$ , which we denote by  $\hat{\vartheta}_{1,0}$  here, is simply  $\hat{\beta}_{1,0}$  divided by  $\bar{w}$ , as the variable  $v_i$  is defined as  $w_i$  divided by  $\bar{w}$ . Finally,  $\text{Cov}(p, v)$  is  $\text{Cov}(p, w)$  divided by  $\bar{w}$  for the same reason.

	Parent 1	Parent 2	$\bar{w}$	$\Delta\bar{p}$	$\hat{\beta}_{1,0}$	$\text{Cov}(p, w)$	$\hat{\vartheta}_{1,0}$	$\text{Cov}(p, v)$
Transition 1	0	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$	1	$\frac{1}{2}$
Transition 2	0	3	$\frac{3}{2}$	$\frac{1}{2}$	3	$\frac{3}{4}$	$\frac{1}{3}$	$\frac{1}{2}$
Transition 3	2	1	$\frac{3}{2}$	$-\frac{1}{6}$	-1	$-\frac{1}{4}$	$-\frac{1}{9}$	$-\frac{1}{6}$
Transition 4	2	3	$\frac{5}{2}$	$\frac{1}{10}$	1	$\frac{1}{4}$	$\frac{1}{25}$	$\frac{1}{10}$
Average	1	2	$\frac{3}{2}$	$\frac{7}{30}$	1	$\frac{1}{4}$	$\frac{71}{225}$	$\frac{7}{30}$

From this table, we can see, as before, that it is not generally true that  $E[\bar{w}]E[\Delta\bar{p}] = E[\hat{\beta}_{1,0}]\text{Var}(p)$ , as

$$\frac{3}{2} \cdot \frac{7}{30} \neq 1 \cdot \frac{1}{4}$$

Also, we can see that it is not generally true that  $E[\bar{w}]E[\Delta\bar{p}] = E[\text{Cov}(p, w)]$ , as

$$\frac{3}{2} \cdot \frac{7}{30} \neq \frac{1}{4}$$

Also, we can see that it is not generally true that  $E[\Delta\bar{p}] = E[\hat{\vartheta}_{1,0}]\text{Var}(p)$ , as

$$\frac{7}{30} \neq \frac{71}{225} \cdot \frac{1}{4}$$

The only thing that is true, is that  $E[\Delta\bar{p}] = E[\text{Cov}(p, v)]$ . This suggests that both the switch from the regression to the covariance form, and the switch from  $w$  to  $v$  are necessary – while regression coefficients relating to  $v$  do not allow for a meaningful interpretation.

### Infinite populations

With ever increasing populations, however, the variance in  $\bar{w}$ ,  $p'$ , and  $\hat{\beta}_{1,0}$  decreases ever more. In the limit of infinitely large populations, the dynamics become deterministic, and the (Generalized) Price equation will also hold again.

To illustrate this, we assume the same model as above, but now not with one parent without the gene, and one with, but with infinitely many parents, while we still assume that there are equally many parents with and without the gene. Under that assumption, the dynamics become deterministic. That implies that  $\bar{w} = \frac{3}{2}$ , which coincides with the expected value for the example when there were only two members in the parent population. It also implies that  $\hat{\beta}_{1,0} = 1$ , which also matches the expected value in the earlier setting, and  $\text{Var}(p)$  is also still  $\frac{1}{4}$ . The only term that changes, is  $\Delta\bar{p}$ ; with deterministic dynamics,  $p' = \frac{1 \cdot 0 + 2 \cdot 1}{1 \cdot 1 + 2 \cdot 1} = \frac{2}{3}$ , which makes  $\Delta\bar{p} = \frac{1}{6}$ .

$$\frac{3}{2} \cdot \frac{1}{6} = 1 \cdot \frac{1}{4}$$

### Notation

With infinitely large populations, it is useful to change the notation. Instead of going over individuals and representing what  $p$ -scores they have, one would have to go over  $p$ -scores and indicate the relative size of the population is that has that  $p$ -score.

To keep it relatively simple, we can think of a finite set of possible  $p$ -scores an individual can have, for instance  $\left\{0, \frac{1}{L}, \frac{2}{L}, \dots, 1\right\}$ . An infinite population can then be characterized by a vector  $x = [x_0, x_1, \dots, x_L]$ , in which  $x_l$  measures the (relative) quantity of individuals with  $p$ -score  $\frac{l}{L}$ . One can assume that these are shares of the total population, in which case  $\sum_{l=0}^L x_l = 1$ . Alternatively, one can imagine that also with infinite populations, the size of the offspring generation as a whole may be larger or smaller than the parent population. This would require the possibility that population vectors do not add up to 1, which would make these numbers not only reflect the relative sizes of the set of individuals with different  $p$ -scores within one population, but also allow for a comparison of the size of the parent population and the offspring population.

The average  $p$ -score in the parent population would then be  $\bar{p} = \frac{\sum_{l=0}^L x_l \frac{l}{L}}{\sum_{l=0}^L x_l}$ . The quantities in the offspring generation could be denoted by  $y = [y_0, y_1, \dots, y_L]$ , and these would be given by  $y_l = x_l \cdot w_l$ . If we assume Model A, as we have above, that means that  $y_l = x_l \cdot \left(\alpha + \beta_{1,0} \frac{l}{L}\right)$ . The average  $p$ -score in the offspring generation is  $\bar{p}' = \frac{\sum_{l=0}^L y_l \frac{l}{L}}{\sum_{l=0}^L y_l}$ , and the variance in  $p$ -score is  $\text{Var}(p) = \sum_{l=1}^L \frac{x_l}{\sum_{k=0}^L x_k} \left(\frac{l}{L}\right)^2 - \left(\sum_{l=1}^L \frac{x_l}{\sum_{k=0}^L x_k} \frac{l}{L}\right)^2$ .

With infinite populations, the randomness disappears, and we are back in a situation where the Generalized Price equation in covariance form holds:

$$\bar{w} \Delta\bar{p} = \text{Cov}(\hat{w}, p)$$

Here we left out the  $E(w\Delta p)$ -term, because with asexual reproduction, this would be 0 always, but also with sexual reproduction, this would be 0 in the limit of infinitely large populations if we assume fair meiosis.

If we then assume Model A, as we have above, then the equation that we had for a parent population, half of which has a  $p$ -score of 0, while the other half has a  $p$ -score of 1, generalizes to an equation that applies to all compositions of the parent population:

$$\bar{w}\Delta\bar{p} = \hat{\beta}_{1,0} \cdot \text{Var}(p)$$

In stochastic models, the (Generalized) Price equation holds for every realization. That does not imply that it holds in expectation. However, if we assume that the population is large enough for the infinite population model to be a good approximation, then the Generalized Price equation will be informative about the dynamics. Misspecification concerns of course still apply.

### Heterozygote advantage revisited

With the relation between modeling and the Price equation sorted out, we can revisit selection with and without heterozygote advantage. This means that we go back to a diploid setting, in which the  $p$ -scores can be 0,  $\frac{1}{2}$ , and 1.

#### Model A

First, we start with the linear model, which is  $w_i = \alpha + \beta_1 p_i + \varepsilon_i$ . We called this Model A at the beginning of this section, and we will also call it Model A here. We furthermore assume an infinite population and fair meiosis. Under those assumptions, the Generalized Price equation for the linear model, applied to the linear model, is

$$\bar{w}\Delta\bar{p} = \hat{\beta}_1 \text{Var}(p)$$

Here,  $\hat{\beta}_1$  is equal to the true  $\beta_1$ , and if we now would like to translate this equation into a criterion for whether or not (or, possibly: when) this gene is selected for, then this is really straightforward: at all frequencies, the gene is selected for if

$$\beta_1 > 0$$

We will refer to this as the rule for selection of a non-social trait with linear fitness effects.

#### Model B

If we take the quadratic model, then the fitness function is  $w_i = \alpha + \beta_1 p_i + \beta_2 p_i^2 + \varepsilon_i$ . In the beginning of this section, we compared Model A to a different Model B, but here we will nonetheless refer to the quadratic model as Model B too, as it will be clear what this refers to here. The Generalized Price equation for the quadratic model, applied to the quadratic model, is

$$\bar{w}\Delta\bar{p} = \hat{\beta}_1 \text{Var}(p) + \hat{\beta}_2 \text{Cov}(p, p^2)$$

Here,  $\hat{\beta}_1$  is equal to the true  $\beta_1$ , and  $\hat{\beta}_2$  is equal to the true  $\beta_2$ . If we now would like to translate this equation into a criterion for when this gene is selected for, then this is still straightforward, but less concise: the gene is selected for if

$$\beta_1 \text{Var}(p) + \beta_2 \text{Cov}(p, p^2) > 0$$

If we now furthermore assume random mating between generations, then at an allele frequency of  $p$ , the shares of individuals with a  $p$ -score of 0,  $\frac{1}{2}$ , and 1 are:  $(1 - p)^2$ ,  $2p(1 - p)$ , and  $p^2$ , respectively. That makes

$$\begin{aligned} \text{Var}(p) &= E[p^2] - E^2[p] = \frac{1}{4} 2p(1 - p) + p^2 - \left(\frac{1}{2} 2p(1 - p) + p^2\right)^2 \\ &= \frac{1}{2} p(1 - p) + p^2 - p^2 = \frac{1}{2} p(1 - p) \end{aligned}$$

and

$$\begin{aligned} \text{Cov}(p, p^2) &= E[p^3] - E[p]E[p^2] = \left(\frac{1}{8} 2p(1 - p) + p^2\right) - p \left(\frac{1}{2} p(1 - p) + p^2\right) \\ &= \frac{1}{4} p + \frac{3}{4} p^2 - p \left(\frac{1}{2} p + \frac{1}{2} p^2\right) = \frac{1}{4} p + \frac{1}{4} p^2 - \frac{1}{2} p^3 = \frac{1}{2} p(1 - p) \left(\frac{1}{2} + p\right) \end{aligned}$$

With random mating, the criterion for when this gene is selected for then becomes

$$\beta_1 + \beta_2 \left(\frac{1}{2} + p\right) > 0$$

Heterozygote advantage would mean that  $\beta_1 > 0$  and  $\beta_2 < 0$ , and if we for instance take  $\beta_1 = 1$  and  $\beta_2 = -1$ , then that simplifies to

$$p < \frac{1}{2}$$

An important feature is that the rule for selection now depends on the population state, as  $\text{Var}(p)$  and  $\text{Cov}(p, p^2)$  are not multiples of each other. This implies that  $\bar{w}\Delta\bar{p}$  may be positive for some, and negative for other population states. With additional assumptions, such as random mating, this reduces to frequency dependence – which is something we should expect with heterozygote advantage. The  $\hat{\beta}_1$  and  $\hat{\beta}_2$ , however, are not dependent on the population state, as they are equal to the true  $\beta_1$  and  $\beta_2$ , and therefore constant. Nice!

### The Generalized Price equation for Model A applied to Model B

The Generalized Price equation for the linear model, applied to the quadratic model, is

$$\bar{w}\Delta\bar{p} = \hat{\beta}_1 \text{Var}(p)$$

with  $\hat{\beta}_1 = \beta_1 + \beta_2 \frac{\text{Cov}(p, p^2)}{\text{Var}(p)}$ . This also gets the direction of selection right for every population state, but here the  $\hat{\beta}_1$  is not an estimator of anything in the quadratic model. As a

symptom of this,  $\hat{\beta}_1$  depends on the population state. With the additional assumption of random mating,  $\frac{\text{Cov}(p, p^2)}{\text{Var}(p)} = \left(\frac{1}{2} + p\right)$ , and therefore, under that assumption,  $\hat{\beta}_1 = \beta_1 + \beta_2 \left(\frac{1}{2} + p\right)$ , which makes the  $\hat{\beta}_1$  vary with  $p$ .

## 6. Discussion

One of the contributions of the Generalized Price equation is that it helps explain the lack of convergence in the debate on the Price equation, as well as the lack of convergence in the debate on the results that are derived with it (which we will come back to in the twin TI discussion paper on the general version of Hamilton's rule). If we draw the positions with somewhat broad strokes, then on the one hand there is a majority position that has confidence in the method of using the original Price equation to derive results, and that believes those results to be correct and meaningful (see for instance [7], [9-21]). On the other hand, there is a minority position that disagrees with the results as well as the method (see for instance [2-5], [22-26]). I claim that the lack of convergence in the debate is the result of looking at everything through the lens of the original Price equation, instead of the generalized version. The original Price equation, which we understand to be the Generalized Price equation for the standard linear model, is mismatched when paired with models that are different, or more general, or with data that do not support the standard linear model. Empirical or theoretical claims that would be correct and meaningful under the standard linear model, or for data that would support the standard linear model, are easily interpreted to also be correct and meaningful in general. This is facilitated by the fact that the original Price equation is an identity, which does not cease to hold when it is combined with models other than the standard linear model, or with data that are generated by other models. If not paired with statistics, the original Price equation does not differentiate between data that support the standard linear model and data that do not, when applied to data – and neither do the other Price-like equations we arrive at when we combine the Generalized Price equation with other statistical models. While most of the papers from the minority position, including mine, are negative, in the sense that they point to where this can be problematic, the Generalized Price equation not only points to the problem, but also gives a solution, as it helps formulate correct and meaningful alternatives to accurately describe the population genetic dynamics, when the way in which fitnesses depend on  $p$ -scores does not fit the standard linear model.

**It cannot be wrong because it is an identity.** The arguments in defense of the way the Price equation is used in the literature typically include the fact that it is an identity (and therefore that it cannot be wrong) as well as references to its generality. In order to counterbalance this, it is worthwhile, first of all, to point out that neither being an identity nor being general is all that special. There are literally infinitely many equations, all of which are identities, and all of which are completely general in the same way that the Price equation is general. In a failed attempt to be funny, I pointed out in an earlier paper [3] that if we take the original Price equation in covariance form, and divide  $\text{Cov}(w, p)$ , not by  $\text{Var}(p)$ , but by the Planck constant  $h$  times the number of times Denmark won the Eurovision Song Contest (denoted by  $DK$ ), then we can rewrite the Price equation as

$$\bar{w}\Delta\bar{p} = \frac{\text{Cov}(w, p)}{h \cdot DK} \cdot h \cdot DK + E(w\Delta p)$$

As a next step, we can define  $\gamma$  as  $\frac{\text{Cov}(w, p)}{h \cdot DK}$ , and write

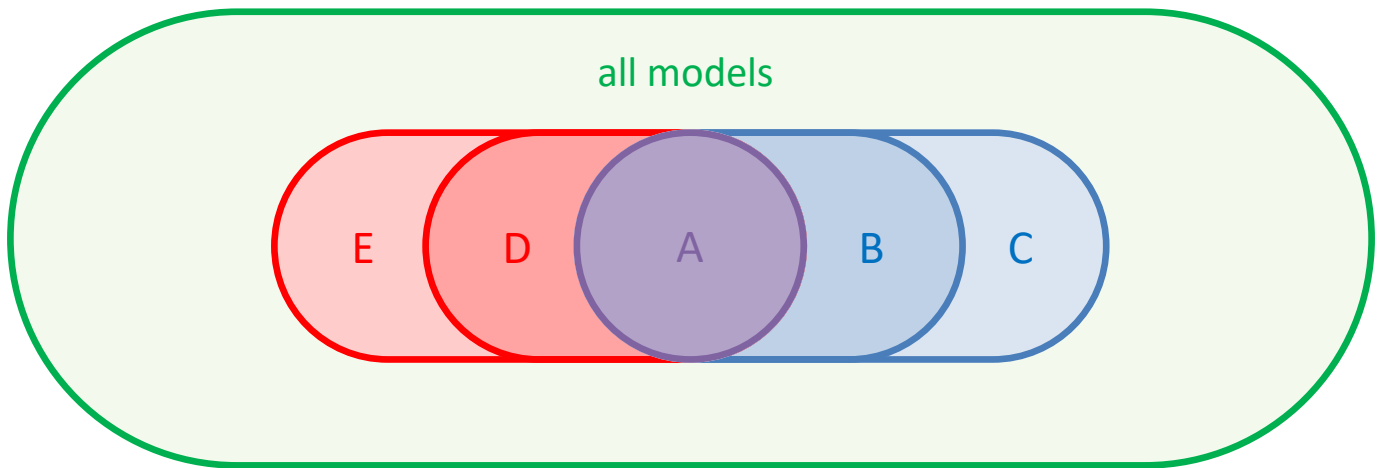
$$\bar{w}\Delta\bar{p} = \gamma \cdot h \cdot DK + E(w\Delta p)$$

This equation is every bit as much an identity as the original Price equation in regression form, and it is also every bit as general. The equation as a whole, and the coefficient  $\gamma$  in particular, however, are clearly devoid of any use or meaning. What this example illustrates, is that it is not enough to be an identity. What we need is an identity with a meaningful interpretation. The original Price equation does have a meaningful interpretation for data that are generated by the standard linear model. We find that the Generalized Price equation, combined with different statistical models, produces other equations, that also are identities and that also are general, and that have a meaningful interpretation for data that are indeed generated by their presupposed statistical models. Having a meaningful interpretation therefore is limited to a *subset* of the models, or of the possible datasets, that the equations can be applied to. When dealing with data, one would have to resort to standard statistics to see if the data can inform us about the model that has generated them. Having a meaningful interpretation then is limited to those datasets that make us decide in favour of the model associated with any specific Price-like equation.

**What we can use the Generalized Price equation for.** In a modeling context, we either directly assume a relation between  $p$ -scores and fitnesses, or we have a model that implies such a relation. For instance, we can assume a model  $w_i = \alpha + \sum_{r=1}^R \beta_r p_i^r$  for some  $R$ , in combination with the assumption of unbiased transmission. Alternatively, we can make modeling assumptions that imply such a relation between  $p$ -scores and fitnesses. The according Price-like equation in regression form then is

$$\bar{w}\Delta\bar{p} = \sum_{r=1}^R \beta_r \text{Cov}(p, p^r)$$

This implies that all terms on the right-hand side are a combination of a model parameter,  $\beta_r$ , that is independent of the population state, and a term that reflects a property of a population state,  $\text{Cov}(p, p^r)$ . The equation then helps state how in infinitely large populations, the change in average  $p$ -score, corrected for the growth or shrinking of the overall population, on the left-hand side, depends on the combination of the current population state, and the model parameters, which we see on the right-hand side. This now describes the population genetic dynamics for all population states, as the model parameters  $\beta_r$  are constants that do not change with the population state. This may or may not be helpful for finding out properties of the model that we are interested in. In Section 4 we have also seen that we can choose richer sets of models too, as long as models in it include a constant and a term that is linear in the  $p$ -score.



**Fig. 3 | Nested models and their Price-like equations.** There are different models, and each model has its own Price-like equation. These Price-like equations are general, in the sense that they can be written for *any* dataset, regardless of the underlying data generating process or, in a theory context, for *any* model. The terms in it, however, only have a meaningful interpretation if the data are generated by a model in the set that the Price-like equation belongs to. In line with the setup in this paper, set A would represent models that are linear in the  $p$ -score ( $w_i = \alpha + \beta_{1,0}p_i$ ). This is the set of models for which the regression coefficient in the original Price equation in regression form has a meaningful interpretation. Set B could consist of models that are quadratic in the  $p$ -score ( $w_i = \alpha + \beta_{1,0}p_i + \beta_{2,0}p_i^2$ ), set C could consist of models that also include a coefficient for the  $p$ -score to the power 3 ( $w_i = \alpha + \beta_{1,0}p_i + \beta_{2,0}p_i^2 + \beta_{3,0}p_i^3$ ), set D models that are linear in the  $p$ -score and the  $q$ -score ( $w_i = \alpha + \beta_{1,0}p_i + \beta_{0,1}q_i$ ), and set E models that include an interaction term between the  $p$ -score and the  $q$ -score ( $w_i = \alpha + \beta_{1,0}p_i + \beta_{0,1}q_i + \beta_{1,1}p_iq_i$ ). The Price-like equations for these different models are all different.

**Dynamic (in)sufficiency revisited.** One possibility for a full model is that the way fitnesses depend on the genes that an individual carries is all there is to it. In that case, a model  $w_i = \alpha + \sum_{r=1}^R \beta_r p_i^r$  maps any parent population state  $p$  to an offspring population state  $p'$ . For simplicity, we assume that this is a deterministic model, or an infinitely large population, which also makes this transition deterministic. (The numbering of the offspring generation, that is, which individual is labeled individual 1, and so on, is irrelevant; all covariances are invariant to permutations of the population vector). We can then iterate this update step, which makes this a dynamically sufficient population dynamics.

One reasonable observation to make here, is that the Price equation itself is not something of which it is useful to describe it as dynamically sufficient or not; dynamic sufficiency is a relevant property of the model it is applied to, and not of the Price equation itself [5]. If, on the other hand, one sees the Price equation as a tool to compute properties of the offspring generation, then it is tempting to describe it as dynamically insufficient. The Price equation gives the new average  $p$ -score, and uses the old  $p$ -score and the covariances  $\text{Cov}(p, p^r)$  pertaining to the first generation to get there. If we then want to repeat this update step, and go from the second population state to the third one, we would need more than the average  $p$ -score in the second state, because now we also need the covariances  $\text{Cov}(p, p^r)$  pertaining to the second generation. These covariances in the second state we can get from the model itself, but we cannot get those from applying the Price equation to the first population state. In that sense, one could be tempted to say that the Price equation is not



dynamically sufficient; once we have reduced what we know about a population state to its average  $p$ -score, we cannot apply it again.

In the twin TI discussion paper on the general version of Hamilton's rule, we will use the Price equation to derive Hamilton's rule, and there, we have a setting in which the  $q$ -score does not reflect the dose of a gene in the individual itself, but in its partner. The model can then describe how the fitness of an individual depends on its  $p$ - and  $q$ -score, and this would produce the  $p$ -scores in the new generation. It would however not generate which individuals are partnered up with whom in the new generation, and therefore it would not identify who has which  $q$ -score in the new generation. In this case a full model would have to include more than the fitness function. In the absence of assumptions about the matching in the new generation, such a model would be dynamically insufficient.

Notice that this is **not** the same as the regression coefficients  $\hat{\beta}_r$ , or  $\hat{\beta}_{k,l}$ , depending on the population state, which is regularly also referred to as dynamical insufficiency. This is incorrect, and really a symptom of misspecification. Describing the dependency of the regression coefficients on the population state as dynamical insufficiency is just not correct, and it can have the effect of not recognizing this as a red flag.

**Summarizing.** The paper in which the Price equation was presented was ambiguous about whether this was meant for statistics or modeling, and it was a bit loose with using terms from statistics and probability theory. That seemed harmless, but it has led to lasting damage in this field. Compared to the derivation of the original Price equation, it takes only a few extra lines of algebra to derive the Generalized Price equation. Behind those extra lines, however, hides an understanding that the original Price equation in regression form is really just one out of a variety of possible equations, and that the terms in it only have meaning if the data for this transition justify the conclusion that they are generated by the linear model. The original Price equation can however be written for *any* change between a parent and an offspring generation, also changes that are very unlikely to have been generated by the linear model that would give the terms in the classical Price equation meaning. For alternative models, one can write alternative Price-like equations, and those are equally general, in the sense that these can be written for any transition too. These other Price-like equations come with the same limitations on the interpretation of the terms in it; if the data do not justify concluding that the model the Price-like equation is build off of is accurate, the terms in it lose their meaning. The Generalized Price equation does however offer a Price-like equation for every model. The classical Price equation is just one of them; the one for the linear model.

# Appendix

## A. Calculations for Example 2.1

$$p = [1,0], w = [2,0] \text{ and } A = P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\bar{w} = \frac{1}{2}(2 + 0) = 1$$

$$\Delta\bar{p} = \bar{p}' - \bar{p} = \frac{1}{2}(1 + 1) - \frac{1}{2}(1 + 0) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$E(w\Delta p) = \frac{1}{n} \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right) = \frac{1}{n} \sum_{i=1}^n \left( \sum_{j=1}^m P_{ij} - w_i p_i \right) = \frac{1}{2}(2) - \frac{1}{2}(2) = 0$$

$$\text{Cov}(w, p) = \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i = \frac{1}{2}(2 \cdot 1 + 0 \cdot 0) - \frac{1}{4}(2 + 0)(1 + 0) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$\text{Var}(p) = \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2 = \frac{1}{2}(1^2 + 0^2) - \frac{1}{4}(1 + 0)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$\beta = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = \frac{\frac{1}{2}}{\frac{1}{4}} = 2$$

## B. Calculations for Example 2.2

$$p = \left[1, \frac{1}{2}\right], w = [2,0] \text{ and } A = P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\bar{w} = \frac{1}{2}(2 + 0) = 1$$

$$\Delta\bar{p} = \bar{p}' - \bar{p} = \frac{1}{2}(1 + 1) - \frac{1}{2}\left(1 + \frac{1}{2}\right) = 1 - \frac{3}{4} = \frac{1}{4}$$

$$E(w\Delta p) = \frac{1}{n} \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right) = \frac{1}{n} \sum_{i=1}^n \left( \sum_{j=1}^m P_{ij} - w_i p_i \right) = \frac{1}{2}(2) - \frac{1}{2}(2) = 0$$

$$\text{Cov}(w, p) = \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i = \frac{1}{2}\left(2 \cdot 1 + 0 \cdot \frac{1}{2}\right) - \frac{1}{4}(2 + 0)\left(1 + \frac{1}{2}\right) = 1 - \frac{3}{4} = \frac{1}{4}$$

$$\text{Var}(p) = \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2 = \frac{1}{2}\left(1^2 + \left(\frac{1}{2}\right)^2\right) - \frac{1}{4}\left(1 + \frac{1}{2}\right)^2 = \frac{5}{8} - \frac{9}{16} = \frac{1}{16}$$

$$\beta = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = \frac{\frac{1}{4}}{\frac{1}{16}} = 4$$

### C. Calculations for Example 3.1

$$p = [1, 0, \frac{1}{2}, 0], w = [\frac{3}{2}, \frac{1}{2}, 1, 1], A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \text{ and } P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

$$\bar{w} = \frac{1}{4} \left( \frac{3}{2} + \frac{1}{2} + 1 + 1 \right) = 1$$

$$\Delta \bar{p} = \bar{p}' - \bar{p} = \frac{1}{4} \left( \frac{1}{2} + 1 + \frac{1}{2} + 0 \right) - \frac{1}{4} \left( 1 + 0 + \frac{1}{2} + 0 \right) = \frac{1}{8}$$

$$\begin{aligned} E(w \Delta p) &= \frac{1}{n} \sum_{i=1}^n w_i \left( \frac{\sum_{j=1}^m P_{ij}}{w_i} - p_i \right) = \frac{1}{n} \sum_{i=1}^n \left( \sum_{j=1}^m P_{ij} - w_i p_i \right) \\ &= \frac{1}{4} \left( \frac{3}{2} - \frac{3}{2} + 0 - 0 + \frac{1}{2} - \frac{1}{2} + 0 - 0 \right) = 0 \end{aligned}$$

$$\begin{aligned} \text{Cov}(w, p) &= \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i \\ &= \frac{1}{4} \left( \frac{3}{2} \cdot 1 + \frac{1}{2} \cdot 0 + 1 \cdot \frac{1}{2} + 1 \cdot 0 \right) - \frac{1}{16} \left( \frac{3}{2} + \frac{1}{2} + 1 + 1 \right) \left( 1 + 0 + \frac{1}{2} + 0 \right) \\ &= \frac{1}{2} - \frac{1}{16} (4) \left( \frac{3}{2} \right) = \frac{1}{2} - \frac{3}{8} = \frac{1}{8} \end{aligned}$$

$$\begin{aligned} \text{Var}(p) &= \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2 = \frac{1}{4} \left( 1^2 + 0^2 + \left( \frac{1}{2} \right)^2 + 0^2 \right) - \frac{1}{16} \left( 1 + 0 + \frac{1}{2} + 0 \right)^2 = \\ &= \frac{1}{4} \cdot \frac{5}{4} - \frac{1}{16} \cdot \frac{9}{4} = \frac{11}{64} \end{aligned}$$

$$\hat{\beta} = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = \frac{\frac{1}{8}}{\frac{11}{64}} = \frac{8}{11}$$

## D. Calculations for Example 3.2

**The stationary distribution.** In order to be able to calculate the frequencies of the different  $p$ -scores in the offspring generation, we first establish that from the parent generation, all 4 parents with a  $p$ -score of  $\frac{1}{2}$  survive, as well as 2 (out of the 4) parents with a  $p$ -score of  $\frac{1}{2}$  (step 1 in Fig. 2). Half of those are female and half of those are male, and therefore we can also think of this as a population where 12 out of 18 survive, 4 females and 4 males with a  $p$ -score of  $\frac{1}{2}$ , and 2 females and 2 males with a  $p$ -score of 1 (step 2).

These are randomly matched, which means that two thirds of the 4 females with a  $p$ -score of  $\frac{1}{2}$  (out of 18 parents) are matched with a male that also has a  $p$ -score of  $\frac{1}{2}$ , and one third is matched with a male that has a  $p$ -score of  $\frac{1}{2}$ , resulting in  $\frac{2}{3} \cdot \frac{4}{18} = \frac{4}{27}$  parent pairs in which both parents have a  $p$ -score of  $\frac{1}{2}$ , and  $\frac{1}{3} \cdot \frac{4}{18} = \frac{2}{27}$  parent pairs with a female with a  $p$ -score of  $\frac{1}{2}$  and a male with a  $p$ -score of 1. Similarly, there will be  $\frac{1}{3} \cdot \frac{2}{18} = \frac{1}{27}$  parent pairs in which both parents have a  $p$ -score of 1, and  $\frac{2}{3} \cdot \frac{2}{18} = \frac{2}{27}$  parent pairs with a female with a  $p$ -score of 1 and a male with a  $p$ -score of  $\frac{1}{2}$  (step 3 and 4).

If both parents have a  $p$ -score of  $\frac{1}{2}$ , 1 out of 4 kids has a  $p$ -score of 0; 2 have a  $p$ -score of  $\frac{1}{2}$ ; and 1 has a  $p$ -score of 1. If one parent has a  $p$ -score of  $\frac{1}{2}$  and the other has a  $p$ -score of 1, then half of their offspring will have a  $p$ -score of  $\frac{1}{2}$ , and the other half will have a  $p$ -score of 1. Finally, if both parents have a  $p$ -score of 1, all of their offspring will have a  $p$ -score of 1 too (step 5).

All parent pairs have 3 kids on average, which then scales this back to the frequencies in the parent population. For example, that results in  $3 \cdot \frac{1}{4} \cdot \frac{4}{27} = \frac{1}{9}$  kids with a  $p$ -score of 0, where 3 is the number of kids per parent pair,  $\frac{1}{4}$  is the proportion of kids with a  $p$ -score of 0 that a parent pair gets when both parents have a  $p$ -score of  $\frac{1}{2}$ , and  $\frac{4}{27}$  measures the number of such parent pairs (step 6).

### The Price equation

Elements of the Price equation for the stationary population state:

$$\bar{w} = \frac{1}{9} \cdot 0 + \frac{4}{9} \cdot \frac{3}{2} + \frac{4}{9} \cdot \frac{1}{2} \cdot \frac{3}{2} = 0 + \frac{6}{9} + \frac{3}{9} = 1$$

$$\Delta \bar{p} = \bar{p}' - \bar{p} = \frac{2}{3} - \frac{2}{3} = 0$$

$$E(w\Delta p) = 0$$

$$\begin{aligned}
\text{Cov}(w, p) &= \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i \\
&= \left( \frac{1}{9} \cdot 0 + \frac{4}{9} \cdot \frac{3}{2} \cdot \frac{1}{2} + \frac{4}{9} \cdot \frac{1}{2} \cdot \frac{3}{2} \cdot 1 \right) - \left( \frac{1}{9} \cdot 0 + \frac{4}{9} \cdot \frac{3}{2} + \frac{4}{9} \cdot \frac{1}{2} \cdot \frac{3}{2} \right) \left( \frac{1}{9} \cdot 0 + \frac{4}{9} \cdot \frac{1}{2} + \frac{4}{9} \cdot 1 \right) \\
&= \frac{6}{9} - 1 \cdot \frac{6}{9} = 0
\end{aligned}$$

$$\begin{aligned}
\text{Var}(p) &= \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2 = \left( \frac{1}{9} \cdot 0^2 + \frac{4}{9} \cdot \left( \frac{1}{2} \right)^2 + \frac{4}{9} \cdot 1^2 \right) - \left( \frac{1}{9} \cdot 0 + \frac{4}{9} \cdot \frac{1}{2} + \frac{4}{9} \cdot 1 \right)^2 = \\
&= \frac{5}{9} - \frac{36}{81} = \frac{1}{9}
\end{aligned}$$

$$\hat{\beta} = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = 0$$

In general, we can assume  $k$  individuals with a  $p$ -score of 0,  $l$  individuals with a  $p$ -score of  $\frac{1}{2}$ , and  $n - k - l$  individuals with a  $p$ -score of 1. The ingredients of the Price equation without noise then become:

$$\bar{w} = \frac{1}{n} \left( k \cdot 0 \cdot \frac{3}{2} + l \cdot 1 \cdot \frac{3}{2} + (n - k - l) \cdot \frac{1}{2} \cdot \frac{3}{2} \right) = \frac{1}{n} \cdot \frac{3}{2} \left( l + (n - k - l) \cdot \frac{1}{2} \right) = \frac{3}{4} \cdot \frac{n - k + l}{n}$$

$$\begin{aligned}
\Delta \bar{p} &= \bar{p}' - \bar{p} = \frac{4}{3} \cdot \frac{1}{n - k + l} \left( k \cdot 0 \cdot 0 \cdot \frac{3}{2} + l \cdot 1 \cdot \frac{1}{2} \cdot \frac{3}{2} + (n - k - l) \cdot \frac{1}{2} \cdot 1 \cdot \frac{3}{2} \right) \\
&\quad - \frac{1}{n} \left( k \cdot 0 + l \cdot \frac{1}{2} + (n - k - l) \cdot 1 \right) = \frac{n - k}{n - k + l} - \frac{n - k - \frac{1}{2}l}{n} \\
&= \frac{n(n - k)}{(n - k + l)n} - \frac{(n - k)^2 + \frac{1}{2}l(n - k) - \frac{1}{2}l^2}{(n - k + l)n} = \frac{k(n - k) - \frac{1}{2}l(n - k) + \frac{1}{2}l^2}{(n - k + l)n} = \frac{k(n - k) - \frac{1}{2}l(n - k - l)}{(n - k + l)n}
\end{aligned}$$

$$E(w \Delta p) = 0$$

$$\begin{aligned}
\text{Cov}(w, p) &= \frac{1}{n} \sum_{i=1}^n w_i p_i - \frac{1}{n^2} \sum_{i=1}^n w_i \sum_{i=1}^n p_i = \frac{1}{n} \sum_{i=1}^n w_i p_i - \bar{w} \bar{p} \\
&= \frac{1}{n} \left( k \cdot 0 \cdot \frac{3}{2} \cdot 0 + l \cdot 1 \cdot \frac{3}{2} \cdot \frac{1}{2} + (n - k - l) \cdot \frac{1}{2} \cdot \frac{3}{2} \cdot 1 \right) - \frac{3}{4} \cdot \frac{n - k + l}{n} \cdot \frac{n - k - \frac{1}{2}l}{n} \\
&= \frac{3}{4} \cdot \frac{n - k}{n} - \frac{3}{4} \cdot \frac{n - k + l}{n} \cdot \frac{n - k - \frac{1}{2}l}{n} = \frac{3}{4} \cdot \left( \frac{n(n - k)}{n^2} - \frac{(n - k)^2 + \frac{1}{2}l(n - k) - \frac{1}{2}l^2}{n^2} \right) \\
&= \frac{3}{4} \cdot \frac{k(n - k) - \frac{1}{2}l(n - k - l)}{n^2}
\end{aligned}$$

$$\begin{aligned}
\text{Var}(p) &= \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2 = \frac{1}{n} \left( k \cdot 0 + l \cdot \left( \frac{1}{2} \right)^2 + (n - k - l) \cdot 1^2 \right) - \left( \frac{n - k - \frac{1}{2}l}{n} \right)^2 = \\
&= \frac{n - k - \frac{3}{4}l}{n} - \left( \frac{n - k - \frac{1}{2}l}{n} \right)^2 = \frac{n^2 - nk - \frac{3}{4}nl}{n^2} - \frac{(n - k)^2 - l(n - k) + \frac{1}{4}l^2}{n^2} = \frac{(k + l)(n - k) - \frac{3}{4}nl - \frac{1}{4}l^2}{n^2}
\end{aligned}$$

$$\hat{\beta} = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = \frac{3}{4} \cdot \frac{k(n - k) - \frac{1}{2}l(n - k - l)}{(k + l)(n - k) - \frac{3}{4}nl - \frac{1}{4}l^2}$$

When there are only individuals with a  $p$ -score of 0, or with a  $p$ -score of  $\frac{1}{2}$ , we have  $k + l = n$ , and the  $\beta$  is  $\frac{kl-0}{nl-\frac{3}{4}nl-\frac{1}{4}l^2} = \frac{kl}{\frac{1}{4}nl-\frac{1}{4}l^2} = \frac{k}{\frac{1}{4}(n-l)} = \frac{k}{\frac{1}{4}k} = 4$ .

When there are only individuals with a  $p$ -score of  $\frac{1}{2}$ , or with a  $p$ -score of 1, we have  $l = 0$ , and the  $\beta$  is  $\frac{k(n-k)-0}{(k+0)(n-k)-0-0} = \frac{k(n-k)}{k(n-k)} = 1$ .

When there are only individuals with a  $p$ -score of 0, or with a  $p$ -score of 1, we have  $k = 0$ , and the  $\beta$  is  $\frac{0-\frac{1}{2}l(n-l)}{ln-\frac{3}{4}nl-\frac{1}{4}l^2} = \frac{-\frac{1}{2}l(n-l)}{\frac{1}{4}nl-\frac{1}{4}l^2} = \frac{-\frac{1}{2}(n-l)}{\frac{1}{4}n-\frac{1}{4}l} = -2$ .

This illustrates the range of possible  $\hat{\beta}$ 's.

Just as a sanity check, if we choose  $k = 1$ ,  $l = 4$  and  $n = 9$ , as in the stationary population state, we do indeed get

$$\bar{w} = \frac{3}{4} \cdot \frac{n-k+l}{n} = \frac{3}{4} \cdot \frac{12}{9} = 1$$

$$\Delta\bar{p} = \frac{k(n-k)-\frac{1}{2}l(n-k-l)}{(n-k+l)n} = \frac{8-8}{12 \cdot 9} = 0$$

$$E(w\Delta p) = 0$$

$$\text{Cov}(w, p) = \frac{3}{4} \cdot \frac{k(n-k)-\frac{1}{2}l(n-k-l)}{n^2} = \frac{3}{4} \cdot \frac{8-8}{9^2} = 0$$

$$\text{Var}(p) = \frac{(k+l)(n-k)-\frac{3}{4}nl-\frac{1}{4}l^2}{n^2} = \frac{5 \cdot 8 - \frac{3}{4} \cdot 9 \cdot 4 - \frac{1}{4} \cdot 4^2}{9^2} = \frac{40 - 27 - 4}{9^2} = \frac{9}{9^2} = \frac{1}{9}$$

$$\hat{\beta} = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = 0$$

### E. Calculations for Example 3.3

The average fitness is just the size of the offspring generation over the size of the parent population;

$$\bar{w} = \frac{m}{n}$$

The change in  $p$ -score between the generations is one half times the difference between the frequency of females in the offspring generation and the frequency of females in the parent population;

$$\Delta\bar{p} = \frac{l + \frac{1}{2}(m-l)}{m} - \frac{k + \frac{1}{2}(n-k)}{n} = \frac{1}{2} \left( \frac{l}{m} - \frac{k}{n} \right)$$

If reproduction happens with equal probabilities on girls and boys, one would expect  $\frac{k}{n}$  to be close to  $\frac{1}{2}$  for large populations.

When calculating the  $E(w\Delta p)$  term, we make use of the fact that all mothers have a  $p$ -score of 1, and therefore their successful gametes must all have  $p$ -score  $\frac{1}{2}$ . This implies that the sum of the  $p$ -scores of their successful gametes must be  $\frac{1}{2}$  times the number of individuals in the offspring generation;  $\sum_{i=1}^k \sum_{j=1}^m P_{ij} = \frac{1}{2}m$ . Also, it implies that  $w_i p_i = w_i$  for all mothers, and therefore  $\sum_{i=1}^k w_i p_i = \sum_{i=1}^k w_i = \frac{1}{2}m$ .

All fathers of daughters must have had a successful gamete with a  $p$ -score of  $\frac{1}{2}$ , while all fathers of sons must have had a successful gamete with  $p$ -score 0. Therefore, the sum of their successful gametes must be  $\frac{1}{2}$  times the total number of daughters;  $\sum_{i=k+1}^n \sum_{j=1}^m P_{ij} = \frac{1}{2}l$ . Also, it implies that  $w_i p_i = \frac{1}{2}w_i$  for all fathers, and therefore  $\sum_{i=k+1}^n w_i p_i = \sum_{i=k+1}^n \frac{1}{2}w_i = \frac{1}{4}m$ .

For mothers and fathers together therefore,  $\sum_{i=1}^n \sum_{j=1}^m P_{ij} = \frac{1}{2}(m+l)$ , and  $\sum_{i=1}^n w_i p_i = \frac{1}{2}m + \frac{1}{4}m = \frac{3}{4}m$ . Therefore

$$E(w\Delta p) = \frac{1}{n} \sum_{i=1}^n \left( \sum_{j=1}^m P_{ij} - w_i p_i \right) = \frac{1}{n} \left( \frac{1}{2}l - \frac{1}{4}m \right) = \frac{1}{2n} \left( l - \frac{1}{2}m \right)$$

When calculating the  $\text{Cov}(w, p)$  term, we use  $\sum_{i=1}^n w_i p_i = \frac{3}{4}m$  again.

$$\begin{aligned} \text{Cov}(w, p) &= \frac{1}{n} \left( \sum_{i=1}^n w_i p_i \right) - \frac{1}{n^2} \left( \sum_{i=1}^k w_i + \sum_{i=k+1}^n w_i \right) \left( \sum_{i=1}^k p_i + \sum_{i=k+1}^n p_i \right) \\ &= \frac{1}{n} \left( \frac{3}{4}m \right) - \frac{1}{n^2} \left( \frac{1}{2}m + \frac{1}{2}m \right) \left( k + \frac{1}{2}(n-k) \right) = \frac{3m}{4n} - \frac{1}{n^2} (m) \frac{1}{2} (n+k) \\ &= \frac{3m}{4n} - \frac{m}{2n} - \frac{mk}{2n^2} = \frac{m}{4n} - \frac{mk}{2n^2} = \frac{m}{2n} \left( \frac{1}{2} - \frac{k}{n} \right) \end{aligned}$$

$$\begin{aligned} \text{Var}(p) &= \frac{1}{n} \sum_{i=1}^n p_i^2 - \frac{1}{n^2} \left( \sum_{i=1}^n p_i \right)^2 = \frac{1}{n} \left( k \cdot 1^2 + (n-k) \left( \frac{1}{2} \right)^2 \right) - \frac{1}{n^2} \left( k + (n-k) \frac{1}{2} \right)^2 \\ &= \frac{1}{n} \cdot \frac{1}{4} (3k+n) - \frac{1}{n^2} \cdot \frac{1}{4} (k+n)^2 = \frac{3k}{4n} + \frac{1}{4} - \left( \frac{1k^2}{4n^2} + \frac{1k}{2n} + \frac{1}{4} \right) = \frac{k}{4n} \left( 1 - \frac{k}{n} \right) \end{aligned}$$

$$\hat{\beta} = \frac{\text{Cov}(w, p)}{\text{Var}(p)} = \frac{\frac{m}{2n} \left( \frac{1}{2} - \frac{k}{n} \right)}{\frac{1k}{4n} \left( 1 - \frac{k}{n} \right)} = \frac{2m \left( \frac{n}{2} - k \right)}{k(n-k)} = \frac{m(n-2k)}{k(n-k)}$$

This implies that  $\hat{\beta} > 0$  if  $k < \frac{1}{2}n$ .

## References

- [1] G. R. Price, 'Selection and Covariance', *Nature*, vol. 227, no. 5257, pp. 520–521, Aug. 1970, doi: 10.1038/227520a0.
- [2] M. van Veelen, 'On the use of the Price equation', *J Theor Biol*, vol. 237, no. 4, 2005, doi: 10.1016/j.jtbi.2005.04.0A.
- [3] M. van Veelen, 'The problem with the Price equation', *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 375, no. 1797, p. 20190355, Apr. 2020, doi: 10.1098/rstb.2019.0355.
- [4] M. van Veelen, J. García, M. W. Sabelis, and M. Egas, 'Call for a return to rigour in models', *Nature*, vol. 467, no. 7316, 2010, doi: 10.1038/467661d.
- [5] M. van Veelen, J. García, M. W. Sabelis, and M. Egas, 'Group selection and inclusive fitness are not equivalent; the Price equation vs. models and statistics', *J Theor Biol*, vol. 299, 2012, doi: 10.1016/j.jtbi.2011.07.025.
- [6] G. R. Price, 'Extension of covariance selection mathematics', *Ann Hum Genet*, vol. 35, no. 4, pp. 485-490., 1972.
- [7] A. Grafen, 'A geometric view of relatedness', *Oxford surveys in evolutionary biology*, vol. 2, no. 2, pp. 28–89, 1985.
- [8] A. Grafen, 'Developments of the Price equation and natural selection under uncertainty', *Proc R Soc Lond B Biol Sci*, vol. 267, no. 1449, pp. 1223–1227, Jun. 2000, doi: 10.1098/rspb.2000.1131.
- [9] P. D. Taylor, 'Evolutionary stability in one-parameter models under weak selection', *Theor Popul Biol*, vol. 36, no. 2, pp. 125–143, Oct. 1989, doi: 10.1016/0040-5809(89)90025-7.
- [10] P. D. Taylor and S. A. Frank, 'How to Make a Kin Selection Model', *J Theor Biol*, vol. 180, no. 1, pp. 27–37, May 1996, doi: 10.1006/jtbi.1996.0075.
- [11] S. A. Frank, *Foundations of social evolution*. Princeton: Princeton University Press, 1998.
- [12] S. Rice, *Evolutionary theory: mathematical and conceptual foundations*. Sinauer Associates, 2004.
- [13] S. Okasha, *Evolution and the levels of selection*. Clarendon Press, 2006.
- [14] A. Grafen, 'Optimization of inclusive fitness', *J Theor Biol*, vol. 238, no. 3, pp. 541–563, Feb. 2006, doi: 10.1016/j.jtbi.2005.06.009.
- [15] A. Gardner, 'The Price equation', *Current Biology*, vol. 18, no. 5, pp. R198–R202, Mar. 2008, doi: 10.1016/j.cub.2008.01.005.
- [16] A. Gardner, S. A. West, and G. Wild, 'The genetical theory of kin selection', *J Evol Biol*, vol. 24, no. 5, pp. 1020–1043, May 2011, doi: 10.1111/j.1420-9101.2011.02236.x.
- [17] S. A. Frank, 'Natural selection. IV. The Price equation\*', *J Evol Biol*, vol. 25, no. 6, pp. 1002–1019, Jun. 2012, doi: 10.1111/j.1420-9101.2012.02498.x.
- [18] F. Rousset, 'Regression, least squares, and the general version of inclusive fitness', *Evolution (N Y)*, vol. 69, no. 11, pp. 2963–2970, Nov. 2015, doi: 10.1111/evo.12791.
- [19] J. A. R. Marshall, *Social evolution and inclusive fitness theory: an introduction*. Princeton University Press, 2015.
- [20] D. C. Queller, 'Fundamental Theorems of Evolution', *Am Nat*, vol. 189, no. 4, pp. 345–353, Apr. 2017, doi: 10.1086/690937.
- [21] V. J. Luque, 'One equation to rule them all: a philosophical analysis of the Price equation', *Biol Philos*, vol. 32, no. 1, pp. 97–125, Jan. 2017, doi: 10.1007/s10539-016-9538-y.



- [22] M. van Veelen, ‘Can Hamilton’s rule be violated?’, *Elife*, vol. 7, Oct. 2018, doi: 10.7554/eLife.41901.
- [23] M. van Veelen, ‘The group selection–inclusive fitness equivalence claim: not true and not relevant’, *Evol Hum Sci*, vol. 2, p. e11, Apr. 2020, doi: 10.1017/ehs.2020.9.
- [24] M. van Veelen, B. Allen, M. Hoffman, B. Simon, and C. Veller, ‘Hamilton’s rule’, *J Theor Biol*, vol. 414, 2017, doi: 10.1016/j.jtbi.2016.08.019.
- [25] B. Allen, M. A. Nowak, and E. O. Wilson, ‘Limitations of inclusive fitness’, *Proceedings of the National Academy of Sciences*, vol. 110, no. 50, pp. 20135–20139, Dec. 2013, doi: 10.1073/pnas.1317588110.
- [26] M. A. Nowak, A. McAvoy, B. Allen, and E. O. Wilson, ‘The general form of Hamilton’s rule makes no predictions and cannot be tested empirically’, *Proceedings of the National Academy of Sciences*, vol. 114, no. 22, pp. 5665–5670, May 2017, doi: 10.1073/pnas.1701805114.