Manuscript accepted for publication in *Philosophical Transactions of the Royal Society B*

The Price equation and the unity of social evolution theory

Jussi Lehtonen
The University of Sydney, Faculty of Science, School of Life and Environmental Sciences, Sydney, 2006 New South Wales, Australia.
E-Mail: jussi.lehtonen@iki.fi (corresponding author)

Abstract

The Price equation has been entangled with social evolution theory from the start. It has been used to derive the most general versions of kin selection theory, and Price himself produced a multilevel equation which provides an alternative formulation of social evolution theory, dividing selection into components between and within groups. In this sense, the Price equation forms a basis for both kin and group selection, so often pitted against each other in the literature. Contextual analysis and the neighbour approach are prominent alternatives for analysing group selection. I discuss these four approaches to social evolution theory and their connections to the Price equation, focusing on their similarities and common mathematical structure. Despite different notations and modelling traditions, all four approaches are ultimately linked by a common set of mathematical components, revealing their underlying unity in a transparent way. The Price equation can similarly be used in the derivation of streamlined, weak selection social evolution modelling methods. These weak selection models are practical and powerful methods for constructing models in evolutionary and behavioural ecology, they can clarify the causal structure of models, and can be easily converted between the four social evolution approaches just like their regression counterparts.

Keywords

Price equation, kin selection, social evolution, group selection, multilevel selection, contextual analysis, neighbour approach

1. The (single-level) Price equation and social evolution theory

From its inception, the price equation has been entangled with social evolution theory. Price was largely motivated by Hamilton's work on altruism [1, P320], wanting to verify the result that to him seemed very grim in the limited way that it seemed to allow 'true' altruism beyond nepotism to evolve [2, P142, 3]. By taking a very different and at the time unique approach [4, 5], Price found Hamilton's results to be correct in their main findings. Furthermore, he found that not only were Hamilton's results correct, but that exactly the same underlying mathematical foundation describes the evolution of altruistic traits and any other naturally selected traits, including spiteful ones [3, P209]. At the level of pure, bare-bones mathematical description of selection, there really seemed to be nothing special about altruism. Nevertheless, his derivation of what is now known as the Price equation has proven valuable for social evolution theory and beyond [6-12]. The Price equation can be presented in a handful of equivalent forms, with the form of equation 1 being perhaps the most common.

$$\bar{w}\Delta\bar{g} = cov(w, g) + E[w\Delta g] \qquad (1)$$

Here w is individual fitness, g is an individual's character value, overbars denote population averages, and Δg denotes change in character value between parent and offspring. While a charitable interpretation might suggest that the Price equation (equation 1) in its generality is all we need to understand social evolution, it seems unlikely that the single-level Price equation alone would have been recognised as a solution to the problem of altruism had it been available prior to the publication of kin selection theory by Hamilton [13-15]. Although both scientists were interested in similar questions, in some ways equation (1) and kin selection theory are very different: Hamilton's aim was to find an explicitly causal explanation for altruism [16-19]. Price's approach, on the other hand, is completely indifferent to the various causal processes that may affect fitness. Equation (1) by itself can, at best, be described as a very 'coarse grained' analysis [20] of altruism. In this case the finer grained analysis (kin selection) came first [13-15], and later merged with the coarse-grained Price equation [16, 21]. The single-level Price equation (1) nevertheless shows that what matters for the evolution of a trait (assuming no transmission bias, i.e. focusing on the first term of equation 1) is the statistical association between the trait and fitness, regardless of causal effects. In particular, with some biological insight, one might see in the statistical formalism of equation (1) that it is not only the direct effect of the trait on the actor's personal fitness that matters for character change – any causal pathway, direct or indirect, that generates positive covariance between the trait and fitness is sufficient. But the Price equation shows this in hindsight, with the benefit of Hamilton's highly intuitive formulations of kin selection.

This is analogous to later debate surrounding kin selection theory. One of the criticisms against kin selection is that kin selection theory is not needed, because the same calculations can always be done with the tools of standard natural selection [22]. Analogously, one might claim that kin selection or multilevel selection are not needed because all the results that could possibly be derived using those methods could be derived using the single-level Price equation. While these claims may be true to some extent, many find the causally more explicit tools of kin selection and multilevel selection helpful.

The aim of this article is to provide a synthesis of four core approaches to social evolution theory, of their relationships to each other, and to the Price equation. Several articles have sought unified perspectives to various aspects of social evolution theory [e.g. 23, 24-27]. The main novel contribution of the current article is to seek a simple unified mathematical foundation for all four central social evolution approaches at their most general level of definition, providing transparency in their relationships, as well as tools for transitioning from one model type to another. This general foundation is then related to weak selection models. As we will see, all social evolution methods are to some extent connected to the Price equation. Price himself would not live to see the influence his work now has. He is said to have had a mixed reaction to his discovery of the equation that now bears his name. On one hand, he was shocked by the implication that at the most fundamental level, altruism doesn't really exist, or at least it is described by exactly the same mathematics as selfishness, spite, or any other trait [3]. But he also saw it as a miracle that he (having not been trained as a biologist) should discover such a fundamental and previously undiscovered equation describing evolution [1, P322-323]. Sadly, Price's life was characterised by internal conflicts of this sort until its untimely end by his own hand [1-3, 28, Harman in this issue].

2. General models: their meaning and their value

In this article we will initially examine four different kinds of models of social evolution at the most general, genic regression level of analysis [16, 25, 29-31]. In other words, we carry out derivations 'as if' we knew the genes for a trait we are interested in and the relevant gene-fitness relationships. Of course we do not really need to know all this to be able to derive the models – we simply need to know certain mathematical facts about regression and covariance. Covariances and regression coefficients in these models are measures of relationships between genetic value (e.g. allele frequency, breeding value) and fitness and they are population statistics as Price emphasised [5], not sample statistics. These relationships exist and obey the mathematical rules of regression regardless of whether someone is around to estimate them or not [18]. The value of such general and somewhat abstract models is one of many highly divisive issues related to social evolution theory. In brief, perhaps the main criticism is that they cannot make testable predictions [32]. The main line of defence is that prediction is not the aim of this type of model, and instead the most general, regression-based formulation forms a robust conceptual and theoretical foundation [29], and provides an organising framework for social evolution theory [18]. In practice, when working with biological questions, we typically use less general and more practical methods [16, 29]. While this criticism and discussion regarding testable predictions has usually revolved around kin selection, it applies equally (to the extent that it is true) to any of the other modelling approaches in this article when formulated using general regression models.

There is no doubt that for many researchers these general models have value, despite their limited value as predictive modelling tools. Their limitations in this regard are the same as those of the Price equation in its most general form: a general model of this kind, based on statistical associations, cannot make predictions without additional assumptions, regardless of whether it is couched in the covariance formalism of the Price equation [4, 5] or the multiple regression formalism of general kin selection models [16]. The models could in principle be applied to any relevant dataset, and model parameters can be calculated from the dataset retrospectively, making the model true almost by definition. But because of this 'retrospective' generality, they provide a very general organising framework for models that make more specific assumptions. In a hypothetical world where all we had to work with was the Price equation and the resulting general models with no further biological or mathematical insight to guide us, they might indeed not be very useful. They are valuable in conjunction with biological insight and with simplifying assumptions that we can make as a consequence. Queller (whom the general model of kin selection is credited to) was in some ways ahead of the critics in writing that the "*real value of this genetic version is that it serves to provide a broader context for inclusive fitness theory*", and noted that in practice it will often be necessary to use less general models [16].

Another strength of general models, and one that has been exploited in previous studies [e.g. 30, 31, 33, 34] is that they provide a natural framework for studying the equivalence of different social evolution methods and their connections. Some articles pointing out limitations of one method or another have in fact compared, say, a general model of multilevel selection to a less general model of kin selection, guaranteeing from the outset that the former will emerge as the winner in a comparison of generality. Again, this was stated succinctly early on in a defence of the equivalence of kin selection and the multilevel Price equation: "*Conclusions to the contrary are due to setting up unfair competitions between the two models*" [33]. Using general models to study equivalences between methods is the approach that will be taken below in section 4.

It has been stated that while general models serve as an organising framework, they are not used to derive more detailed evolutionary models [18]. While this may be a fair description of the usual state of affairs, it seems a little restrictive. If any model that makes more specific assumptions than the general models counts as a 'more detailed evolutionary model', then it is possible – and sometimes useful – to derive these more specific results using the general models as a starting point. It can be helpful to roughly think of the general models as a template, or outer layer that contains the less general variants in a nested structure. For example, (partial) regression coefficients which appear in general models and (partial) derivatives which appear in streamlined methods are analogous and coincide under certain conditions [25, 31, 35] (see section 5). When these conditions are fulfilled, one can see the structure of a weak selection model based on differentiation directly from the structure of a general model based on regression, and 'derive' less general formulations from the general framework by taking advantage of the equivalences. This could be visualised as a move down in a structure where the most general models are at the top, and less general ones further down (Figure 1). As long as one is careful with the simplifying assumptions that are being made, it is often possible to move down but not back up once the model has been 'degeneralised'. For example, a weak selection model cannot be said to contain a genic regression model as a special case even if the opposite is true (although the structure of a less general model may give hints as to what a more general model might look like). One can also move horizontally to derive equivalence results between different models at similar levels of generality. I have previously used this approach to first derive an equivalence result and transformation rule between kin selection and multilevel selection at the most general, genic regression level (a horizontal move in figure 1), and then relate this equivalence result to the less general weak selection level of models (a vertical move) [31]. We will revisit model transformations like this in sections 4-5.

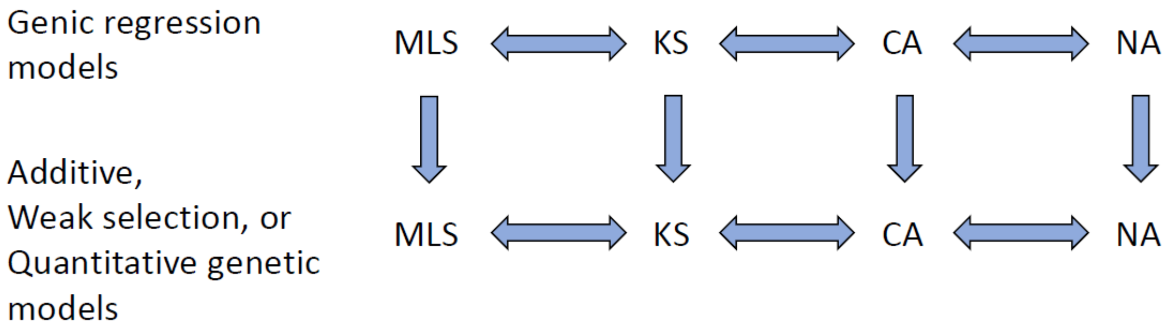

Figure 1. The hierarchy of social evolution methods. MLS=multilevel Price equation; KS=kin selection; CA=contextual analysis; NA=neighbour approach. The most general forms of models (genic regression models [16]) form a robust foundation for social evolution theory. The different approaches are mathematically equivalent when compared at the same level of generality (horizontal arrows). The general forms can be thought to contain less general forms as special cases when certain conditions are fulfilled [e.g. 33, 34], but the opposite is not true: it is possible to move downwards from more to less general, but not vice-versa once simplifying assumptions have been made. Note that all possible pairs of models on each row could be connected by a double-headed arrow (e.g. MLS ↔ CA). For clarity and simplicity, only arrows between adjacent pairs are included in the figure, thus connecting all pairs of models via a series of double-headed arrows.

So, to form a unified and mutually comparable basis, all the methods will be initially defined as genic regression models. While contextual analysis (section 3c), for example, is commonly applied as a phenotypic method [36], there is no fundamental obstacle preventing the same general methodology from being specified at a genic level of analysis. Only the nature of the independent variables in equations (6) and (7) change when we switch between genotype and phenotype, while the logical structure of the underlying model remains the same and its relationship to other methods is exposed more clearly. The difference between phenotypic and genotypic methods becomes even smaller if we accept the perspective of Rice [7]: "*alleles and genotypes can be thought of as particular kinds of phenotypes*". Hence, taking such a genic perspective [in the sense of 16] in all the modelling alternatives described below, we are free to compare alternative approaches to social evolution theory while avoiding the inevitable differences that arise from unequal starting points, without straying too far from the original meaning of the models. If one chooses to focus on differences between methods, one will always find differences. Here the focus is on similarities, in the spirit of the claim that debates between methods are often largely about historical differences between approaches, not deep logical differences [37].

3. Four approaches to social evolution theory

In this section we will examine four alternative approaches to social evolution theory in their most general forms. We begin with the two that are most intimately tied to the Price equation: the multilevel Price equation and kin selection. Although the last two methods, contextual analysis and the neighbour approach are not as closely linked to it, we will make use of the Price equation when converting them from a model of fitness to a model of evolutionary change. The aim is to examine all models at equal levels of generality while using the same notation (Table 1) as far as possible to expose their similarities with maximum transparency.

Table 1. Notation and definitions.

Notation	**Definition**	**Rationale and examples**
g	Individual character value	A characteristic of an individual; in this article this denotes a genic value, e.g. allele frequency or breeding value
G	Whole-group mean character value	Mean value of g over the entire group that the individual belongs to
$\tilde{G}$	Other-only mean character value	Mean value of g over the group that the individual belongs to, excluding the focal individual itself
$\bar{g}$	Population mean character	Mean value of g over the entire population
w	Individual fitness	Fitness of an individual
W	Whole-group mean fitness	Mean value of w over the entire group that the individual belongs to
$\bar{w}$	Population mean fitness	Mean value of w over the entire population
$b, \tilde{b}$	Benefit	The 'benefit' of Hamilton's rule, for whole-group and other-only models respectively
$c, \tilde{c}$	Cost	The 'cost' of Hamilton's rule, for whole-group and other-only models respectively
$R, \tilde{R}$	Relatedness	Coefficient of relatedness of Hamilton's rule, for whole-group and other-only models respectively
β_{ab}	Simple regression coefficient of a on b	For example, β_{Gg} is the regression coefficient of group mean character value on individual character value
$\beta_{ab.c}$	Partial regression coefficient of a on b, with c held constant	For example, $\beta_{wg.\tilde{G}}$ is the partial regression coefficient of individual fitness on individual character value, with mean character value of everyone else in the group held constant

a) Group selection 1: The many faces of the multilevel Price equation

The multilevel version of the Price equation (equation 2 – henceforth termed MLS) is one of the central theoretical approaches to social evolution theory. It was briefly presented in a follow-up publication by Price himself [5], and later revisited in various forms by several other authors [e.g. 1 P318, 33, 38, 39, 40]. Price's original form for the multilevel equation is

$$\bar{w}\Delta\bar{g} = \underbrace{cov(W,G)}_{\textit{between-groups}} + \underbrace{E[cov_k(w,g)]}_{\textit{within-groups}} \qquad (2)$$

where the index k indicates that the covariances within the expectation term are group-specific covariances, and the expectation is subsequently taken over all groups. If groups differ in size, group-size weighted expectations and covariances must be used (see appendix).

Where the single-level Price equation can be thought of as a very 'coarse-grained' [20] analysis of social evolution (see section 1), the MLS equation represents a more fine-grained analysis by splitting selection and evolutionary change into two components. It is called a multilevel selection model because it handles selection in a hierarchical manner: evolutionary change is split into two components that are commonly thought to correspond to selection between groups (i.e. group selection), and to selection within groups. Within social evolution theory, the MLS equation is perhaps best known as a modern formulation of group selection that is not at odds with individual level selection. One of the reasons group adaptation is considered the exception rather than the norm in evolution is that in the absence of special conditions, a trait that has arisen purely via selection between groups would tend to be undermined by selection between individuals, within groups [41-43]. The multilevel Price equation accounts for this by including selection within groups alongside selection between groups. Of course, this does not imply that conditions where group adaptation can arise are never fulfilled. For example, the evolution of multicellularity is closely connected to the concept of group adaptation [44], and if we consider multicellular organisms as group adapted groups of cells, then group adaptation has taken place relatively often.

Above I wrote that the two components are 'commonly thought to' correspond to selection between and within groups. The reason for the non-committal phrasing is that simply applying the MLS equation to some set of data (real or hypothetical) does not guarantee that its two components will correspond to selection between and within groups in a causal and biologically meaningful sense [10]. Perhaps the simplest way to see this is to imagine a dataset for which we know the value of a trait and fitness for every individual in a population with no group structure. Despite the lack of group structure, we can assign the individuals to arbitrary, discrete groups if we wish to do so, and compute the between- and within-group components of equation 2 arising from this arbitrary structure. But because we started with a population without group structure, we know from the outset that there is no biologically meaningful selection between groups going on even if the between-group component takes a non-zero value. This example does not mean that the MLS equation is not useful. It simply illustrates that being able to compute values for the two components of equation 2 for some dataset or model does not imply that the mathematical components are meaningful for that particular scenario. To be biologically meaningful, it is necessary that the scenario being studied has biologically meaningful groups. It is fairly common for models to have this kind of discrete group structure: many models in social evolution make use of the infinite island model [45], which entails a natural structure that can often be analysed using the MLS equation. However, although the existence of such groups is a necessary requirement for a sensible interpretation of the MLS equation, interpretation should be done with caution even in the presence of group structure. George Williams emphasised the difference between genuine group adaptation on one hand and fortuitous group benefits that are really by-products of individual level adaptations on the other ([41], see also e.g. [10] for discussion). For example, if good eyesight increases the fitness of an individual (without detrimental effects on other group members) it will also increase average fitness of the group that the individual belongs to: there is a benefit to group fitness, but the group benefit arises as a statistical summation of the effects of individual adaptations. This statistical summation would similarly appear in the between-groups component of the multilevel Price equation. Whether this is a problem depends on what one expects of the equation. If there is variation in eyesight between groups, then it is arguably reasonable to speak of selection between groups and the MLS equation may correctly identify this group component of selection. But one must be cautious in drawing the conclusion that eyesight is a genuine group adaptation. Improved eyesight would evolve whether group structure is present or not.

Price originally derived Equation 2 in a ‘downwards’ fashion, by starting from the between-group level, and working towards the lower, within-groups level [5]. However, in this section I will present a very brief and simple ‘upwards’ derivation of the multilevel selection equation, starting from the individual level and decomposing selection into components between and within groups. Both derivations lead to mathematically equivalent decompositions, and both have advantages. One very useful insight from the original downwards derivation is that transmission bias (or property change in the words of Price) on one level can often be interpreted as selection on a lower level [5]. In other words, selection within groups distorts the outcome from what it would be if selection only took place between groups, causing a kind of transmission bias when viewed at the group level. On the other hand, an advantage of the upwards derivation is that it makes transparent the fact that both components of the multilevel Price equation can be written as covariances over all individuals in the population (as emphasized by Heisler and Damuth [46]), despite the between-groups component looking like a covariance taken only over groups in equation 2. A second advantage of the upwards derivation is that it entails no reason to assume absence of intergroup migration [40], which was an assumption that Price made in his original formulation [5]. The derivation here is indebted to those of Wade [40], Queller [33], and Bijma & Wade [47] although it is not identical to any of them.

We begin with the standard, individual level Price equation with no transmission bias (i.e. the first term of equation 1): $\bar{w}\Delta\bar{g} = cov(w, g)$, where the covariance is taken over the entire population.

Now, if all individuals in the population belong to non-overlapping groups with mean group fitness W, each individual’s fitness can be decomposed into two components – the group mean and individual deviation from the group mean (Δw):

$w = W + (w - W) = W + \Delta w.$

Substituting this into the first term of equation 1 we get

$\bar{w}\Delta\bar{g} = cov(w, g) = cov(W + \Delta w, g)$

Making use of the distributive property of covariance, we can write this as

$$\bar{w}\Delta\bar{g} = \underbrace{cov(W, g)}_{\textit{between-groups}} + \underbrace{cov(\Delta w, g)}_{\textit{within-groups}} \qquad (3)$$

It can be shown that both terms of equation 3 are mathematically equivalent with those of equation 2 (see appendix for derivations). Equations 2 and 3 can both be illuminating. The salient question in the historical group selection debate could be phrased in two ways: Are traits of *individuals* adaptations for the benefit of the group, possibly at the cost of individual interests (the perspective taken, at least in part, by Williams [41])? Or alternatively, can we consider traits of *groups* adaptations for the benefit of the group (a viewpoint taken by e.g. Gardner & Grafen [42])? Equations 2 and 3 show that these partially contrasting questions are in some sense the same. The first term of equation 2 is the statistical association between group fitness and group mean trait value (a common way to define a group trait). The first term of equation 3 is the statistical association between group mean fitness and individual trait value. The equivalence of these terms indicates that from the perspective of the multilevel Price equation, the two above ways of describing group adaptation are identical(it must be conceded, however, that this is only a partial resolution of the contrast described above, because statistical associations are not necessarily indicators of causal relationships – see section 5b for some further problems associated with causal interpretation). There are in fact three ways of writing the first covariance: $cov(W, G) = cov(W, g) = cov(w, G)$ (see appendix). We will return to the last form in section 4. Care must be taken in that for the last two forms, the covariance is computed over all individuals, whereas in the first form (which appears in the 'classic' equation - equation 2) the covariance is computed over all groups and must be weighted by group size if they are unequal in size [5] – no such group size weighting is needed for the last two alternatives.

For the second term we already have two equivalent forms, $E[cov_k(w, g)] = cov(\Delta w, g)$. Two further alternatives are $cov(w, \Delta g) = cov(\Delta w, \Delta g)$. The last form is from Queller ([33], see also [30, 31]). Again, the expectation in equation 2 requires weighting by group size if group sizes are unequal, but this is not necessary in the alternative forms.

b) Kin selection

William Hamilton's theoretical work on the genetical evolution of social behaviour [13-15] (later widely known as 'kin selection', a term actually coined by Maynard Smith [48]) was published well before the Price equation, and served as a major motivator for Price's work (see section 1). As the two scientists corresponded prior to the publication of Price's 1970 article [4], Hamilton came to see the Price equation as a cleaner and more appropriate way to formulate his central findings on social evolution compared to his original rather cumbersome derivation [1 P175]. Hamilton published a Price equation-based rederivation of his central result immediately following the publication of the Price equation itself. In fact the story goes that this was the result of a collusion between the two – Hamilton pressured *Nature* [1] into either publishing Price's paper [4] or risk losing Hamilton's own contribution [21], in a move that seems unlikely by today's standards. The derivation I present here nevertheless owes the most to Queller [16] who derived this now widely known 'general' model of kin selection – general in the sense that minimal assumptions are needed in its derivation.

The starting point for this model is a multiple regression equation for fitness:

$$w = \alpha + \beta_{wg.\tilde{G}} g + \beta_{w\tilde{G}.g} \tilde{G} + \varepsilon \tag{4}$$

This is a direct fitness [49] or neighbour-modulated fitness [14] approach to kin selection, where g represents each individual's genes for the trait being analysed, while $\tilde{G}$ is the average of other individuals that the focal individual's fitness is affected by. $\beta_{wg.\tilde{G}}$ and $\beta_{w\tilde{G}.g}$ are partial regression coefficients of w on g controlling for $\tilde{G}$, and of w on $\tilde{G}$ controlling for g respectively. α and ε represent the intercept and residuals of the regression. Hamilton's rule can be derived by substituting equation 4 into the first term of the Price equation (equation 1) (note that dropping the second term amounts to an assumption of no transmission bias):

$$\bar{w}\Delta\bar{g} = cov(w, g) = \beta_{wg.\tilde{G}} var(g) + \beta_{w\tilde{G}.g} cov\left(\tilde{G}, g\right)$$

A key point here is that because g is a predictor in equation 4, it must be uncorrelated with ε according to least square theory (and with the constant α) [16]. Division by $var(g)$ yields

$$\frac{\bar{w}\Delta\bar{g}}{var(g)} = \beta_{wg.\tilde{G}} + \beta_{w\tilde{G}.g} \frac{cov(\tilde{G}, g)}{var(g)}$$

Here $\frac{cov(\tilde{G},g)}{var(g)} = \beta_{\tilde{G}g}$ is a regression coefficient of relatedness $\tilde{R}$ [29, 50]. If we now denote $\beta_{wg.\tilde{G}} = -\tilde{c}$ and $\beta_{w\tilde{G}.g} = \tilde{b}$ (corresponding to costs and benefits—note the usual sign convention with $-\tilde{c}$), we recover a version of Hamilton's rule [14]: the trait is positively selected for if

$$\frac{\bar{w}\Delta\bar{g}}{var(g)} = -\tilde{c} + \tilde{b}\tilde{R} > 0 \tag{5}$$

This general version of Hamilton's rule gives a cost-benefit condition for gene frequency increase, where costs, benefits and relatedness are defined as regression coefficients. This formulation of kin selection has recently been discussed at length elsewhere [18, 29, 30]. Here I wish to draw attention to one point that will become important later: The group average g-values in equation 4 can be defined in two ways. The value for the focal individual can be either included or excluded from the mean. The latter procedure was followed in the derivation above, in line with Queller [16], but both options are frequently used in kin selection models, and in the notation of this paper (Table 1) we could have equally well used G instead of $\tilde{G}$ in the derivation. These two alternative (but formally equivalent in the domain where both are valid) formulations of kin selection models are often termed 'whole-group' or 'other-only' models. Other-only models correspond to the classical view of kin selection [14], but whole-group models can offer mathematical convenience for certain biological questions. A few details regarding other-only and whole-group models are worth noting. For one, other-only models are in some sense more general: a whole-group relatedness coefficient can only be calculated if individuals are structured into discrete and non-overlapping groups, while other-only relatedness also applies to a more continuously structured population [51]. Second, whole-group relatedness can be thought of as average relatedness to the entire group including oneself [50], or as expected relatedness between two randomly chosen group members with replacement [52]. Consequently, because relatedness to self is one, and because the focal individual forms a larger fraction of smaller groups, whole-group relatedness is group-size dependent. Third, the choice of formulating a model as a whole-group or other-only model typically changes all three central components (b, c, R). In the regression model above this arises because it changes the way the partial regression coefficients are computed. In weak selection models making use of partial derivatives (section 5) this arises because it changes the way partial derivatives are computed. The distinction between other-only and whole-group models will also be central in section 4.

This article focuses on the similarities and common mathematical components of different methods. Therefore we set aside some central causal and philosophical characteristics of kin selection that distinguish it from other methods discussed in this paper. In particular, here we focus on direct or neighbour-modulated fitness instead of inclusive fitness [14, 53]. This is not to downplay the importance of the inclusive fitness perspective, which is valuable in its own right [18, 53-55] – the reason for this choice of focus is simply that because the inclusive fitness perspective is unique to kin selection, commonalities between methods are easier to see using the direct fitness perspective.

c) Group selection 2: Contextual analysis

The origins of contextual analysis lie in the social sciences [56], from which it was adapted for use in evolutionary studies [10, 36, 46, 57-59]. Contextual analysis does not fundamentally rely on the Price equation. In its simplest form, contextual analysis is not necessarily concerned with evolutionary change, and simply makes use of a multiple regression model of individual fitness as the dependent variable, and an individual character and a contextual character as independent variables [46]. Contrasting contextual analysis with the MLS equation (equations 2-3) as a model of social evolution, the former can be simply a model of trait-fitness relationship, whereas the latter is always a model of evolutionary change.

In contextual analysis, one independent variable is always individual character, while the second predictor can in principle be any kind of group character. However, in evolutionary applications the typical group character is the mean character value of the social group. The idea is that because the model works with *partial* regression coefficients, they are intended to isolate the effects of group character and individual character on individual fitness. Contextual analysis is conceptually quite a different way to quantify group selection, compared to the multilevel Price equation. Contextual analysis deals exclusively with individual fitness, and analyses the effect of group character on fitness of individuals. The group selection component in the multilevel Price equation (equations 2-3) on the other hand can be written purely in terms of group fitness, hence it is intuitive to interpret it as that component of evolutionary change that is due to the effect of a trait on group fitness. Neither of the components of contextual analysis alone have this interpretation.

In the notation of this article (Table 1), the regression equation underlying contextual analysis is [58]

$$w = \alpha + \beta_{wg.G} g + \beta_{wG.g} G + \varepsilon \tag{6}$$

Equation (6) can then be converted to one of evolutionary change using the Price equation, analogous to the previous section:

$$\bar{w}\Delta\bar{g} = cov(w,g) = \beta_{wg.G} var(g) + \beta_{wG.g} cov(G,g) = \beta_{wg.G} var(g) + \beta_{wG.g} var(G) \tag{7}$$

d) Group selection 3: The neighbour approach

The neighbour approach is mathematically quite similar to contextual analysis, but despite the apparently minimal differences there are justifications for considering it as an alternative formulation of group selection [10, 51, 58, 60]. Here I will only briefly introduce the mathematical formulation. The neighbour approach can be described as a variant of contextual analysis, where the second independent variable is the mean trait value of everyone but the focal individual, instead of the mean of the entire group [58]:

$$w = \alpha + \beta_{wg.\tilde{G}} g + \beta_{w\tilde{G}.g} \tilde{G} + \varepsilon \tag{8}$$

As with contextual analysis, equation (8) can then be converted to one of evolutionary change using the Price equation:

$$\bar{w}\Delta\bar{g} = cov(w,g) = \beta_{wg.\tilde{G}} var(g) + \beta_{w\tilde{G}.g} cov\left(\tilde{G},g\right) \tag{9}$$

4. A common foundation for the four approaches

It has been stated repeatedly that the methods of section 3 are formally equivalent [10, 18, 26, 30, 33, 58, 61] in terms of evolutionary change, while partitioning selection in different ways. Although true, this may not be obvious or intuitively clear. It can be helpful to show explicitly how they are mathematically equivalent – that is, to find a unifying set of mathematical building blocks that underlies all the above methods. There are two advantages to such a unified view. Firstly, it makes the equivalence of methods completely transparent even with little mathematical training. Second, to the extent that the different formulations can be likened to causal decompositions of selection [61], a unified mathematical language makes it simple to switch from one causal perspective to another, independent of the methodology under which the model was first derived [31]. The benefits of the first point are clearest under the abstract and general model definitions described above. The benefits of the second point are more relevant under concrete models of specific biological scenarios, where the model components (e.g. c, b, and R) gain a more detailed biological meaning.

Many may have already noted similarities among the regression equations in previous sections. In particular, the regression equations underlying kin selection, contextual analysis, and the neighbour approach are very similar. The resemblance between kin selection and contextual analysis in particular is widely recognised [31, 36, 61]. Here I will describe the relationship between contextual analysis, the neighbour approach, and kin selection in more detail, and then relate all three to the multilevel Price equation. Again, an important point in this analysis is that we consider the different methods at the same level of generality (see section 2).

Consider first the regression equation underlying the neighbour approach (equation 8), where the independent variables are individual character value and the mean character value of everyone but the focal individual. Comparing this to the definition of ‘other-only’ kin selection models in section 3a reveals an immediate and obvious link between the two methods. Recall that in other-only kin selection models the independent variables are again individual character value and the mean character value of everyone but the focal individual. Hence the regression model underlying other-only kin selection models is identical to that underlying the neighbour approach when defined at equivalent levels of analysis. Similarly, recall that in an evolutionary context, the independent variables in contextual analysis are typically individual character value and mean character value of the entire group, including the focal individual. Comparing this formulation of contextual analysis to whole-group kin selection models we note that the independent variables are again identical.

Hence we can conclude that the regression equation foundation of other-only kin selection models is identical to that underlying the neighbour approach when defined at equivalent levels of analysis. Similarly, we can conclude that the foundation of whole-group kin selection regression models is identical to that underlying a typical contextual analysis model. Of course, the underlying mathematical similarity doesn’t mean that these approaches are identical in every way. For example, in kin selection models the focus is often on relatedness, which may not be explicitly specified in other types of models (although as we will see, relatedness is implicitly present in all four approaches). Second, and perhaps more importantly, the kin selection models discussed in this article are direct fitness [49] or neighbour-modulated fitness [14] models, whereas kin selection is perhaps better known for the inclusive fitness perspective it affords [14, 53]. A further complication is that kin selection and group selection models are often seen as competing alternatives [48, 62], while both neighbour approach and contextual analysis (which we have just noted to be all but identical to kin selection) have been described as possible ways of modelling group selection [58, 60, 63]. It is not too surprising then that similarities in the foundations of models are easily overlooked in the face of differences in interpretation and terminology.

Nevertheless, we are now in a position to choose a common set of model components which we can use to write the foundations of different modelling approaches in a unified language to make their connections clearer. The *b-c-R* notation we employ is adapted from kin selection, in part because of its simplicity, and partly due to its sufficiency for our purposes: all four approaches to social evolution described above can be written using these mathematical components, without making any additional simplifying assumptions. Keeping in mind that we need to differentiate between other-only and whole-group models, the aim is now to write all of the four methodologies above using only b, c, R (corresponding to regression coefficients $\beta_{wG.g}$, -$\beta_{wg.G}$ and β_{Gg}) or their other-only equivalents $\tilde{b}$, $\tilde{c}$ and $\tilde{R}$ (corresponding to $\beta_{w\tilde{G}.g}$, -$\beta_{wg.\tilde{G}}$ and $\beta_{\tilde{G}g}$).

For contextual analysis and the neighbour approach this is very straightforward. As noted above, the regression equations underlying other-only kin selection models (equation 4) and the neighbour approach (equation 8) are identical, and (using the *b*,*c*,*R* notation of Table 1) can both be written as

$$w = \alpha - \tilde{c}g + \tilde{b}\tilde{G} + \varepsilon \qquad (10)$$

which can be converted to an equation for evolutionary change using the Price equation:

$$\bar{w}\Delta\bar{g} = cov(w, g) = -\tilde{c}\ var(g) + \tilde{b}\ cov(\tilde{G}, g) \qquad (11)$$

Similar equations corresponding to contextual analysis and other-only kin selection are:

$$w = \alpha - cg + bG + \varepsilon \qquad (12)$$

and

$$\bar{w}\Delta\bar{g} = cov(w, g) = -c\ var(g) + b\ cov(G, g) = -c\ var(g) + b\ var(G) \qquad (13)$$

Until now there is no difference between the other-only kin selection versus neighbour approach models on one hand, and whole-group kin selection versus contextual analysis models on the other. Further division of equations (11) and (13) by $var(g)$ yields the typical presentation of Hamilton’s rule written in terms of either other-only or whole-group relatedness:

$$\frac{\bar{w}\Delta\bar{g}}{var(g)} = -\tilde{c} + \tilde{b}\frac{cov(\tilde{G},g)}{var(g)} = -\tilde{c} + \tilde{b}\tilde{R} \tag{14}$$

$$\frac{\bar{w}\Delta\bar{g}}{var(g)} = -c + b\frac{var(G)}{var(g)} = -c + bR \tag{15}$$

So although relatedness is generally not explicitly used in contextual analysis or neighbour approach models, it is in fact implicitly contained in the regression equations, just as it is in the general kin selection models. In terms of the underlying logic of the models, there is very little difference between kin selection, contextual analysis and the neighbour approach. The similarity between kin selection and contextual analysis has been discussed previously in reference [36], and the relationship of other-only and whole-group models to contextual analysis and the neighbour approach has been noted in reference [31].

Moreover, if individuals are in groups that are all equal in size (n), simple relationships can be found between the whole-group and other-only model components [31, 50, 58]: $R = \frac{1+(n-1)\tilde{R}}{n}$, $b = \frac{n}{n-1}\tilde{b}$ and $c = \tilde{c} + \frac{1}{n-1}\tilde{b}$. Therefore relatively simple transitions between any of the aforementioned methods are possible if we are willing to accept the assumption of equal group sizes.

This still leaves open the question of the exact relationship of these methods to the multilevel Price equation. The associations among the other methods were simplified by their basis in regression model, whereas the multilevel Price equation is quite a different looking covariance equation, without an obvious relation to the b, c, and R components. However, if we maintain the same, genic regression level of models, the multilevel Price equation can also be written purely in terms of the whole-group kin selection components, making no simplifying assumptions and allowing arbitrary variation in group sizes:

$$\frac{\bar{w}\Delta\bar{g}}{var(g)} = \underbrace{(b-c)R}_{\textit{between-groups}} \qquad \underbrace{-c(1-R)}_{\textit{within-groups}} \tag{16}$$

Readers are referred to box 2 of ref. [31] for a brief derivation, and the supplemental material of ref. [31] for a more general treatment. The key to the derivation of equation 16 is noting that the MLS equation can be written entirely in terms of individual fitness (section 3a): $\bar{w}\Delta\bar{g} = cov(w,G) + cov(w,g-G)$. Substituting the regression equation 12 into this form of the MLS equation obtains equation 16. If all groups are equal in sizes, we can make use of the relationships between whole-group and other-only model components described above and write the multilevel Price equation in terms of the latter:

$$\frac{\bar{w}\Delta\bar{g}}{var(g)} = \underbrace{\left(\tilde{b}-\tilde{c}\right)\left(\frac{1}{n}+\frac{(n-1)\tilde{R}}{n}\right)}_{\textit{between-groups}} \qquad \underbrace{-\frac{\tilde{b}+(n-1)\tilde{c}}{n}(1-\tilde{R})}_{\textit{within-groups}} \tag{17}$$

When comparing equations 2-3 and 16-17, note that the latter two have been divided by $var(g)$ for clarity. This is a notational preference, and the equations are mathematically equivalent. In equations 2-3 the variance is subsumed in the covariances on the right hand side. Note also that R in equation 16 is whole-group relatedness, and hence takes on a minimum value of $1/n$ even in groups of unrelated individuals. In equation 17, on the other hand, $\tilde{R}$ is other-only relatedness, and $1/n$ is included as a separate component. Thus both equations indicate that the between-groups component can take on non-zero values even if group members are not related to each other, particularly in small groups.

The overall outcome of this section then is that any of the four methods of sections 3a-3d can often be converted to any other. They all have a similar foundation which can be written in terms of a common set of regression coefficients. Although typical applications of, say kin selection and contextual analysis might differ [36], this section also shows that the differences are generally not due to any inherent limitations or strengths of particular models, and more due to the different ways in which researchers choose to apply them.

Table 2. Equations for evolutionary change under alternative approaches to social evolution theory, and relationships between model components using the c-b-R notation (see table 1).

Other only kin selection	$\frac{\bar{w}\Delta\bar{g}}{var(g)} = -\tilde{c} + \tilde{b}\tilde{R}$
Neighbour approach	$\bar{w}\Delta\bar{g} = -\tilde{c}\ var(g) + \tilde{b}\ cov(\tilde{G}, g)$ or $\frac{\bar{w}\Delta\bar{g}}{var(g)} = -\tilde{c} + \tilde{b}\tilde{R}$
Whole-group kin selection	$\frac{\bar{w}\Delta\bar{g}}{var(g)} = -c + bR$
Contextual analysis	$\bar{w}\Delta\bar{g} = -c\ var(g) + b\ var(G)$ or $\frac{\bar{w}\Delta\bar{g}}{var(g)} = -c + bR$
Multilevel Price equation	$\frac{\bar{w}\Delta\bar{g}}{var(g)} = \underbrace{(b-c)R}_{between\text{-}groups} \quad \underbrace{-c(1-R)}_{within\text{-}groups}$
Relatedness	$\tilde{R} = \frac{cov(\tilde{G},g)}{var(g)} = \beta_{\tilde{G}g}$ $R = \frac{cov(G,g)}{var(g)} = \frac{var(G)}{var(g)} = \beta_{Gg}$
Relationship between whole-group and other-only model components (*n*=group size)	$R = \frac{1+(n-1)\tilde{R}}{n}$ $b = \frac{n}{n-1}\tilde{b}$ $c = \tilde{c} + \frac{1}{n-1}\tilde{b}$

Note that whether we retain $var(g)$ and $\bar{w}$ on the left or right side of the equations for evolutionary change is somewhat arbitrary and depends on notational preferences and the requirements of the question we are studying. If we are mainly interested in the direction of change, it may be clearest to retain these factors on the left side (both are non-negative and hence do not influence direction), while moving them to the right side yields rates of change, regardless of which approach we are using [37].

5. Weak selection models, additive models

a) Regression coefficients versus derivatives

While genic regression models have the advantage of being very general and as such providing robust structure to social evolution theory, they are by themselves not practical for many modelling applications. Gardner et al discuss the distinction between "*the general theory of kin selection that forms the foundations of social evolution vs. the streamlined kin selection methodologies that are used to solve specific problems*" [29]. To some extent a similar division into general and streamlined methods can be applied to all social evolution methods described above in their general form.

For a modeller in evolutionary or behavioural ecology, perhaps the most practical methods are weak selection models (more precisely, δ-weak selection in the terminology of Wild and Traulsen [64]) at the interface of game theory and social evolution theory [8]. Powerful and flexible kin selection modelling techniques under weak selection are presented in detail elsewhere [19, 35, 49, 65], allowing for complications such as class structured models. The aim here is not to present the full methodology with all its extensions, but instead to focus on its connection to the general models described above, and on how the four social evolution approaches described in previous sections can be viewed from a weak selection perspective.

Streamlined weak selection methods can be thought to approximate partial regression coefficients with partial derivatives [25, 35], and it is helpful to have a visual understanding of why and when this works. Partial regression coefficients tell us something about the effect of one independent variable on the dependent variable while holding the other independent variables constant, regardless of how much variation in present in the population. Partial derivatives on the other hand tell us about the effect of *small variation* in one variable on a function, with other variables held constant. There is therefore a clear intuitive analogy between partial regression coefficients and partial derivatives. Visualising some aspects of this connection nevertheless becomes much easier if we consider fitness as a non-linear function of one trait and corresponding simple regression coefficients and derivatives, illustrated in Figure 2. As phenotypic variation reduces, the relevant portion of the phenotype-fitness relationship becomes closer and closer to a straight line, the slope of which is given by the derivative of the fitness function. On the other hand, when we fit a regression model to data generated from a straight line, the regression coefficient is exactly the slope of this line. Hence, under small phenotypic variation (δ-weak selection [64]), derivatives and regression coefficients coincide. A two-variable equivalent with partial regression coefficients and partial derivatives is analogous, but more difficult to visualise: it would consist of a fitness surface in 3-dimensional coordinates, approximated by a plane (a higher-dimensional analogue of a straight line). The plane can be defined using partial regression coefficients, or partial derivatives, and again under small variation the two coincide. Replacing (partial) regression coefficients with (partial) derivatives, the latter can be computed using only the population mean value instead of the full phenotypic distribution, taking advantage of the powerful methods of calculus [8].

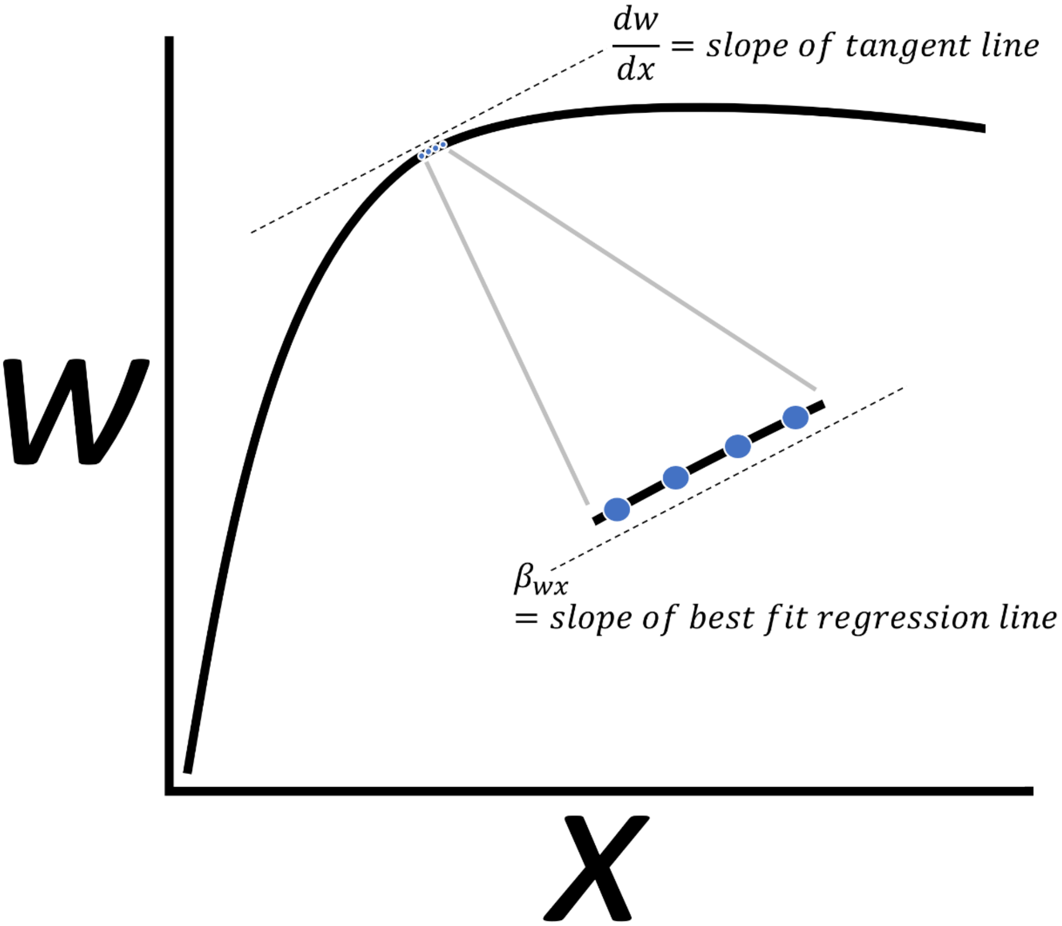

Figure 2. A simplified visualisation of approximating a regression coefficient with a derivative. The curve represents a hypothetical phenotype-fitness relationship, here for clarity assumed to be deterministic so that phenotype determines fitness exactly. If there is very little phenotypic variation in the population at any given time (δ-weak selection), the phenotype-fitness relationship is very close to a straight line for any set of phenotypes simultaneously undergoing selection (the zoomed in section, with blue circles representing individuals). The derivative gives the slope of the tangent line to the curve. A regression coefficient gives the slope of a straight line that best fits the phenotype-fitness data points. As phenotypic variation decreases (the blue circles get closer together), the difference between β_{wx} and $\frac{dw}{dx}$ becomes smaller and smaller. If there is wide phenotypic variation, the derivative can be a very poor approximation of the regression coefficient due to the non-linearity of the phenotype-fitness relationship – in such a case, regression will give the correct outcome while differentiation will not. An analogous explanation applies to partial regression coefficients and partial derivatives but is more difficult to visualise: the curve would be replaced by a surface and the straight line with a plane, both embedded in three-dimensional space.

Figure 2 depicts a phenotype-fitness relationship, whereas all the preceding equations are concerned with genotype-fitness relationships. There are alternative ways to reconcile these views. First, suppose we are mainly concerned with ultimate evolutionary equilibria that arise over very long timescales. In this case, when a wide range of potential mutations is considered, theory suggests that the long-term stable state can be characterised at the phenotypic level [66]. Genetic constraints that temporarily prevent a phenotypic maximum from being reached are eventually broken by altered genetic systems (e.g. variation in scale of mutation effects or dominance), until the final equilibrium is reached. This justifies the use of phenotypic differentiation methods when characterising evolutionary equilibria. Alternatively (or to complement a static analysis of equilibria), if we want to approximate evolutionary trajectories, we can make the simplifying assumption of additive genetic effects and small variation at any given time [8], which implies that phenotype maps onto genotype in a simple way even outside of equilibria. Though obviously a simplification, this may not be too far from reality: mutations on quantitative traits are expected to often have small phenotypic effects, particularly in polygenic traits, where the combined small effects of multiple genes determine the value of a continuous phenotypic trait [29, see 67 for a review of recent empirical evidence]. Either way, phenotypic weak selection models arguably expose selection and the explanation for the appearance of 'design' in a very transparent way. It is a separate question whether genetic constraints can prevent the suggested equilibrium form being reached in the short term.

For a more mathematical perspective of the connection between derivatives and regression coefficients we can take advantage of Taylor polynomials [68] in conjunction with the Price equation. As long as the fitness function satisfies certain mathematical conditions (namely, it is analytic), it can be approximated by a Taylor polynomial. For social evolution models where we need to account for both individual character and that of the social environment, we can approximate fitness using a first order multivariable Taylor polynomial [see reference 8 for details]. In either case, the Price equation can be applied to the Taylor polynomial, yielding an expression for evolutionary change in terms of derivatives. To further convince ourselves that partial derivatives correspond to partial regression coefficients, we could use the first order multivariable Taylor polynomial as an approximate, linear function for fitness and explicitly calculate the partial regression coefficients of previous sections corresponding to the linear approximation (using e.g. the formulas in box 4 of reference [29]).

Whichever way we choose to approach the question, the outcome is that the partial regression coefficients of previous sections have partial derivative counterparts:

$\beta_{wg.G} = -c \approx \frac{\partial w}{\partial g}$ and $\beta_{wG.g} = b \approx \frac{\partial w}{\partial G}$ for whole-group models, and similar for their other-only equivalents. And given that these partial regression coefficients appear in all four social evolution approaches of sections 3-4, we can write weak selection approximations for all of them by replacing the partial regression coefficients with their partial derivative counterpart. The point of this is that even though the weak selection optimality approach is typically applied to kin selection models [36], the difference is again one of historical differences in approach, rather than one relating to fundamental differences in logic [37]. In many cases we could interpret a weak selection social evolution model from any perspective we choose for a particular model and biological question, and arguably, in some cases such a change of viewpoints can be useful.

b) Causal interpretation in regression, additive, and weak selection models

Causality is a slippery concept, and one that has been a source of debate in evolutionary biology for several decades [e.g. 54, 69, 70, 71]. The concept of causality and how it connects with evolutionary theory raises many questions. On a very general scale, we can ask whether adaptive evolution, as described by mathematical evolutionary theory is a causal process in the first place [72]. When examining more specific evolutionary scenarios, we may want to consider what are valid and useful ways of breaking the model into finer grained causal structures [e.g. kin selection versus multilevel Price equation: 61]. Fairly recently, there have been great strides forwards in causal modelling in general [e.g. 73, 74], and these developments provide a rigorous framework in which to tackle such questions. There has also been much interest in the validity of causal interpretation of regression coefficients, particularly in Hamilton's rule [25, 75]. Here I would like to emphasise the perhaps underappreciated counterpoint that weak selection models and associated derivatives and partial derivatives can be a powerful aid for fine-grained causal analysis of natural selection that can supplement causal interpretation using other methods.

What do weak selection models permit in terms of causal interpretation that additive models or regression models do not? An additive model makes the strong assumption that fitness is mediated by additive effects of phenotype. A regression model, on the other hand, 'forces' a nonlinear model into additive components by fitting a model of best linear fit to the data (in the sense of minimising the sum of squares). Both methods have their place and can be valuable, but in some sense, both are a compromise. The assumption of additivity in trait-fitness relationships is known to commonly be broken – additivity is likely the exception rather than the rule in nature. Regression gets around this by extracting statistical summaries from the data that superficially seem additive. But as a downside, because they force a non-additive model into an additive mould, the regression coefficients themselves may not have a clear causal meaning unless the underlying causal model is additive itself [75].

Derivatives and partial derivatives fall somewhere in between, making a third sort of compromise. Differentiation too can extract a type of additive effect from a non-additive model, but in a very different way than regression. It is an effect that is valid only in a specific domain of small variation (Figure 2). But within this limited domain it can represent the real causal structure defined by a potentially non-linear model of fitness.

There is something that may seem puzzling about this. Weak selection methods are typically used to analyse models constructed by researchers themselves. If the causal effects in the model are entirely specified by the modeller, how can differentiation and some sort of causal analysis of the derivatives yield any additional understanding (particularly since, if anything, differentiation *removes* information because all constant terms have the same derivative)? One reason for this is that mere differentiation can be thought of as shorthand for steps that are implicitly taking place behind the scene: first, we derive an additive, weak selection model for fitness by computing a Taylor polynomial for a potentially non-linear and frequency-dependent fitness function, and then apply the Price equation to the resulting approximate additive model of fitness [8], thus obtaining an approximate model for evolutionary change. So, if we break down the process into steps, what looks like differentiation of the fitness function is in fact something relatively complex and by no means obvious. We have gone from a static description of fitness to a model that tracks change of phenotype over time under weak selection. From this perspective, it no longer seems very puzzling that differentiation reveals aspects of causal structure that may not be obvious when we write down the initial fitness function.

Another reason for differentiation being a potentially useful aid in causal analysis is that weak selection methods based on differentiation (e.g. game theory of continuous traits, quantitative genetics under weak selection, adaptive dynamics, the direct fitness approach to kin selection, all of which lead to fairly similar equations [8, 35, 76-80]) are often used to model complex scenarios which can incorporate rich ecological detail. This means that the resulting function for fitness can be quite complex, and the potentially simpler expressions that arise from differentiation may be easier to interpret. The fitness function may include multiple components and complicated nonlinear interactions which can be difficult to disentangle intuitively. Linearising the fitness function using a Taylor approximation can separate the causal components neatly into additive components that are easier to interpret.

An early example of using differentiation to clarify the causal structure of an evolutionary model is Taylor's 1981 analysis of sex ratio models [81]. Since Hamilton's publication of the local mate competition sex ratio model in 1967 [82], there had been much debate and confusion on what it is that really drives the evolution of skewed sex ratios in this model – again, despite the fact that these causal processes are ultimately completely specified by the modeller. Taylor differentiated a simple equation for fitness, and gave a causal interpretation to the partial derivatives that arise from this model, greatly clarifying our understanding of the question using this fairly simple procedure. Therefore using differentiation and δ-weak selection models as a type of causal analysis is not new, even if it has not necessarily been described using those words. But δ-weak selection could potentially be a powerful aid in the ongoing discussion of the causal structure of evolutionary theory. A thorough investigation of these models as an aid in causal analysis, and of the limitations of such an analysis would also help bring philosophical and biological aspects of causal analysis closer to each other. Philosophical contributions often focus on additive or regression models [e.g. 72, 75], while δ-weak selection is one of the most important modelling methods in evolutionary biology, at least in social evolution theory [29].

In evolutionary models, it is typically possible to decompose the total model into component causes in many different ways [19]. From this perspective, kin selection is one kind of causal analysis [19], while other approaches to social evolution can be thought of as alternative causal perspectives [61] (although as we have seen, direct fitness kin selection models, contextual analysis, and the neighbour approach are very similar). The mathematical terms of different approaches can carry different causal meanings, some more appropriate to particular biological questions than others. Mathematical translation rules between different social evolution approaches are then in an informal sense simultaneously translations between different causal perspectives that can be applied under regression, additive and weak selection models. I have previously argued that although perhaps easiest to initially derive using direct fitness kin selection formalism [35], models of gamete evolution under gamete competition and gamete limitation [31], as well as local mate competition sex allocation models [83] can both benefit from an MLS interpretation. It is not that the MLS perspective is in some absolute sense the correct interpretation for these questions, but rather that in both models there are ecological factors that map nicely onto the between-groups and within-groups components of the MLS equation. In models of gamete evolution, variation in the total fertilisation success of a spawning group (modelled with a ‘fertilisation function’ [84]) corresponds to between-groups selection, while competition between spawners (gamete or sperm competition [85]) within each group corresponds to within-groups selection. This clarifies the relationship of early group selectionist models of gamete evolution [86, 87] to later individual selection models [88]. Similarly, models of sex allocation with local mate competition have occasionally been given a group selection interpretation [89, 90]. From an MLS perspective, the total reproductive output of a local reproductive group corresponds to between-groups selection which can select for a female-biased sex ratio, while the within-groups component mirrors standard Fisherian selection [91] for a 50:50 sex ratio [83]. These examples simply illustrate the ease with which one can analyse a single model from multiple perspectives under detailed evolutionary models, while table 2 shows how different social evolution approaches relate to each other under a more general framework.

6. Conclusion

The main point of this article has been to bring a unified perspective to alterative (and often, seemingly competing) approaches to social evolution theory. We began with the standard, ‘single-level’ Price equation, noting that under a broad interpretation it can accommodate the evolution of altruism, spite, or any other kind of trait typically investigated using social evolution theory. But it does not do this by providing a causal explanation for the evolution of these traits, but rather by being completely causally agnostic: any process that creates positive covariance between a trait and fitness entails selection for that trait, no matter what and how complicated the intermediate causal pathways are. Depending on one’s perspective, this causal indifference can be seen as a positive or negative feature of the Price equation.

Examining four more ‘fine-grained’ approaches to social evolution we found that under very broad conditions, they can all be written using a unified set of mathematical components, allowing easy translations from one methodology to another. This unification can be derived under the most general, genic regression models. But because regression coefficients can be identified with derivatives under weak selection, this unification also applies to the very practical and streamlined weak selection models. A similar argument works for additive models. From this perspective, the differences between alternative approaches are minimal, while each still retains specific strengths as modelling methods and causal interpretations of evolution.

Finally, I have argued that weak selection models can be a powerful aid in causal analysis. Typically, weak selection models are used to estimate evolutionary equilibria and evolutionary trajectories. However, the derivatives that make up such a model are often amenable to causal interpretation. Such causal interpretation has proven helpful in clarifying causal relationships in e.g. sex ratio models. Because the translation rules established in this article apply to weak selection models, we are free to view a weak selection model from the alternative causal perspectives that different social evolution approaches afford.

Funding

Jussi Lehtonen was funded by an Australian Research Council Discovery Early Career Research Award (project number DE180100526) from the Australian Government.

Acknowledgements

I am grateful to David Queller and Richard Michod for helpful and constructive comments on the manuscript.

References

[1] Hamilton, W. D. 1996 *Narrow Roads of Gene Land: The Collected Papers of W.D. Hamilton Volume 1: Evolution of Social Behaviour*, W.H. Freeman/Spektrum.
[2] Segerstråle, U. 2013 *Nature's Oracle: The Life and Work of W.D.Hamilton*, OUP Oxford.
[3] Harman, O. 2010 *The Price of Altruism: George Price and the Search for the Origins of Kindness*, W. W. Norton.
[4] Price, G. R. 1970 Selection and covariance. *Nature* **227**, 520-521.
[5] Price, G. R. 1972 Extension of covariance selection mathematics. *Annals of human genetics* **35**, 485-490.
[6] Queller, D. C. 2017 Fundamental Theorems of Evolution. *The American Naturalist* **189**, 345-353. (DOI:doi:10.1086/690937).
[7] Rice, S. H. 2004 *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sunderland, Massachusetts, Macmillan Education.
[8] Lehtonen, J. 2018 The Price Equation, Gradient Dynamics, and Continuous Trait Game Theory. *The American Naturalist* **191**, 146-153. (DOI:10.1086/694891).
[9] Luque, V. J. 2017 One equation to rule them all: a philosophical analysis of the Price equation. *Biology & Philosophy* **32**, 97-125. (DOI:10.1007/s10539-016-9538-y).
[10] Okasha, S. 2006 *Evolution and the Levels of Selection*. Oxford, Oxford University Press.
[11] Fox, J. W. 2006 Using the Price Equation to partition the effects of biodiversity loss on ecosystem function. *Ecology* **87**, 2687-2696.
[12] Helanterä, H. & Uller, T. 2010 The Price equation and extended inheritance. *Philosophy & Theory in Biology* **2**.
[13] Hamilton, W. D. 1963 The Evolution of Altruistic Behavior. *The American Naturalist* **97**, 354-356. (DOI:10.2307/2458473).
[14] Hamilton, W. D. 1964 Genetical evolution of social behaviour I. *Journal of Theoretical Biology* **7**, 1-16.
[15] Hamilton, W. D. 1964 Genetical evolution of social behaviour 2. *Journal of Theoretical Biology* **7**, 17-52.
[16] Queller, D. C. 1992 A general model for kin selection. *Evolution* **46**, 376-380. (DOI:10.2307/2409858).
[17] Queller, D. C. 2011 Expanded social fitness and Hamilton's rule for kin, kith, and kind. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 10792-10799. (DOI:10.1073/pnas.1100298108).
[18] Birch, J. 2017 *The Philosophy of Social Evolution*, Oxford University Press.
[19] Frank, S. A. 2013 Natural selection. VII. History and interpretation of kin selection theory. *Journal of Evolutionary Biology* **26**, 1151-1184.

[20] Birch, J. 2017 The inclusive fitness controversy: finding a way forward. *Royal Society Open Science* **4**, 170335. (DOI:10.1098/rsos.170335).
[21] Hamilton, W. D. 1970 Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**, 1218-1220.
[22] Nowak, M. A., Tarnita, C. E. & Wilson, E. O. 2010 The evolution of eusociality. *Nature* **466**, 1057-1062.
[23] Michod, R. E. & Hamilton, W. D. 1980 Coefficients of relatedness in sociobiology. *Nature* **288**, 694-697.
[24] Nowak, M. A. 2006 Five rules for the evolution of cooperation. *Science* **314**, 1560-1563.
[25] Birch, J. & Okasha, S. 2015 Kin Selection and Its Critics. *Bioscience* **65**, 22-32. (DOI:10.1093/biosci/biu196).
[26] Marshall, J. A. R. 2011 Group selection and kin selection: formally equivalent approaches. *Trends in Ecology & Evolution* **26**, 325-332. (DOI:10.1016/j.tree.2011.04.008).
[27] Michod, R. & Sanderson, M. 1985 Behavioural structure and the evolution of cooperation. In *Evolution: Essays in Honor of John Maynard Smith* (eds. P. J. Greenwood, P. H. Harvey & M. Slatkin), pp. 95-105, CUP Archive.
[28] Frank, S. A. 1995 George Price's contributions to evolutionary genetics. *Journal of Theoretical Biology* **175**, 373-388. (DOI:http://dx.doi.org/10.1006/jtbi.1995.0148).
[29] Gardner, A., West, S. A. & Wild, G. 2011 The genetical theory of kin selection. *Journal of Evolutionary Biology* **24**, 1020-1043. (DOI:10.1111/j.1420-9101.2011.02236.x).
[30] Marshall, J. A. R. 2015 *Social Evolution and Inclusive Fitness Theory: An Introduction*, Princeton University Press.
[31] Lehtonen, J. 2016 Multilevel Selection in Kin Selection Language. *Trends in Ecology & Evolution* **31**, 752-762. (DOI:10.1016/j.tree.2016.07.006).
[32] Nowak, M. A., McAvoy, A., Allen, B. & Wilson, E. O. 2017 The general form of Hamilton's rule makes no predictions and cannot be tested empirically. *Proceedings of the National Academy of Sciences*. (DOI:10.1073/pnas.1701805114).
[33] Queller, D. C. 1992 Quantitative Genetics, Inclusive Fitness, and Group Selection. *The American Naturalist* **139**, 540-558.
[34] Birch, J. & Marshall, J. A. R. 2014 Queller's Separation Condition Explained and Defended. *The American Naturalist* **184**, 531-540. (DOI:10.1086/677750).
[35] Taylor, P. D. & Frank, S. A. 1996 How to make a kin selection model. *Journal of Theoretical Biology* **180**, 27-37.
[36] Goodnight, C. 2013 On multilevel selection and kin selection: contextual analysis meets direct fitness. *Evolution* **67**, 1539-1548. (DOI:10.1111/j.1558-5646.2012.01821.x).
[37] Queller, D. 2016 Kin selection and its discontents. *Philosophy of Science* **83**, 861-872. (DOI:doi:10.1086/687870).
[38] Hamilton, W. D. 1975 Innate social aptitudes of man: an approach from evolutionary genetics. In *Biosocial anthropology* (pp. 133-155.
[39] Wilson, D. S. 1975 A theory of group selection. *Proceedings of the National Academy of Sciences* **72**, 143-146.
[40] Wade, M. J. 1985 Soft Selection, Hard Selection, Kin Selection, and Group Selection. *The American Naturalist* **125**, 61-73. (DOI:10.2307/2461608).
[41] Williams, G. C. 1966 *Adaptation and natural selection: a critique of some current evolutionary thoughts*. Princeton, New Jersey, Princeton University Press.
[42] Gardner, A. & Grafen, A. 2009 Capturing the superorganism: a formal theory of group adaptation. *Journal of Evolutionary Biology* **22**, 659-671. (DOI:10.1111/j.1420-9101.2008.01681.x).
[43] Dawkins, R. 1976 *The Selfish Gene*. Oxford, Oxford University Press.
[44] Shelton, D. E. & Michod, R. E. 2014 Group Selection and Group Adaptation During a Major Evolutionary Transition: Insights from the Evolution of Multicellularity in the Volvocine Algae. *Biol Theory* **9**, 452-469. (DOI:10.1007/s13752-014-0159-x).

[45] Wright, S. 1931 Evolution in Mendelian populations. *Genetics* **16**, 97-159.
[46] Heisler, I. L. & Damuth, J. 1987 A Method for Analyzing Selection in Hierarchically Structured Populations. *The American Naturalist* **130**, 582-602. (DOI:10.2307/2461705).
[47] Bijma, P. & Wade, M. J. 2008 The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *Journal of Evolutionary Biology* **21**, 1175-1188. (DOI:10.1111/j.1420-9101.2008.01550.x).
[48] Maynard Smith, J. 1964 Group Selection and Kin Selection. *Nature* **201**, 1145-1147.
[49] Taylor, P. D., Wild, G. & Gardner, A. 2007 Direct fitness or inclusive fitness: how shall we model kin selection? *Journal of Evolutionary Biology* **20**, 301-309. (DOI:10.1111/j.1420-9101.2006.01196.x).
[50] Pepper, J. W. 2000 Relatedness in trait group models of social evolution. *Journal of Theoretical Biology* **206**, 355-368.
[51] Godfrey-Smith, P. 2008 Varieties of population structure and the levels of selection. *The British Journal for the Philosophy of Science* **59**, 25-50.
[52] Day, T. & Taylor, P. D. 1997 Hamilton's rule meets the Hamiltonian: kin selection on dynamic characters. *Proceedings of the Royal Society B: Biological Sciences* **264**, 639-644. (DOI:10.1098/rspb.1997.0090).
[53] Taylor, P., Wild, G. & Gardner, A. 2007 Direct fitness or inclusive fitness: how shall we model kin selection? *Journal of Evolutionary Biology* **20**, 301-309.
[54] Maynard Smith, J. 1987 How to model evolution. In *The Latest on the Best: Essays on Evolution and Optimality* (ed. J. Dupré), pp. 119-131, MIT Press.
[55] Fromhage, L. & Jennions, M. D. 2019 The strategic reference gene: an organismal theory of inclusive fitness. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20190459. (DOI:doi:10.1098/rspb.2019.0459).
[56] Boyd, L. & Iversen, G. 1979 *Contextual Analysis: Concepts and Statistical Techniques*, Wadsworth Pub Co.
[57] Goodnight, C. J., Schwartz, J. M. & Stevens, L. 1992 Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *American Naturalist*, 743-761.
[58] Okasha, S. 2004 Multilevel selection and the partitioning of covariance: a comparison of three approaches. *Evolution* **58**, 486-494.
[59] Damuth, J. & Heisler, I. L. 1988 Alternative formulations of multilevel selection. *Biol Philos* **3**, 407-430.
[60] Nunney, L. 1985 Group selection, altruism, and structured-deme models. *American Naturalist*, 212-230.
[61] Okasha, S. 2016 The relation between kin and multilevel selection: an approach using causal graphs. *The British Journal for the Philosophy of Science* **67**, 435-470.
[62] Kramer, J. & Meunier, J. 2016 Kin and multilevel selection in social evolution: a never-ending controversy? *F1000Research* **5**, F1000 Faculty Rev-1776. (DOI:10.12688/f1000research.8018.1).
[63] Goodnight, C. J. 2015 Multilevel selection theory and evidence: a critique of Gardner, 2015. *Journal of Evolutionary Biology* **28**, 1734-1746. (DOI:10.1111/jeb.12685).
[64] Wild, G. & Traulsen, A. 2007 The different limits of weak selection and the evolutionary dynamics of finite populations. *Journal of Theoretical Biology* **247**, 382-390. (DOI:http://dx.doi.org/10.1016/j.jtbi.2007.03.015).
[65] Frank, S. A. 1998 *Foundations of social evolution*. Princeton, New Jersey, Princeton University Press.
[66] Hammerstein, P. 1996 Darwinian adaptation, population genetics and the streetcar theory of evolution. *Journal of Mathematical Biology* **34**, 511-532. (DOI:10.1007/bf02409748).
[67] Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A. & Yang, J. 2017 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101**, 5-22. (DOI:10.1016/j.ajhg.2017.06.005).
[68] Weisstein, E. W. 2002 *CRC Concise Encyclopedia of Mathematics, Second Edition*. London, CRC Press.

[69] Sober, E. & Lewontin, R. C. 1982 Artifact, cause and genic selection. *Philosophy of science* **49**, 157-180.
[70] Walsh, D. M. 2010 Not a Sure Thing: Fitness, Probability, and Causation. *Philosophy of Science* **77**, 147-171. (DOI:10.1086/651320).
[71] Otsuka, J., Turner, T., Allen, C. & Lloyd, E. A. 2011 Why the Causal View of Fitness Survives. *Philosophy of Science* **78**, 209-224. (DOI:10.1086/659219).
[72] Otsuka, J. 2016 Causal Foundations of Evolutionary Genetics. *The British Journal for the Philosophy of Science* **67**, 247-269. (DOI:10.1093/bjps/axu039).
[73] Pearl, J. 2000 *Causality: models, reasoning and inference*, Cambridge Univ Press.
[74] Spirtes, P., Glymour, C. N. & Scheines, R. 2000 *Causation, prediction, and search*, MIT press.
[75] Okasha, S. & Martens, J. 2016 The causal meaning of Hamilton's rule. *Royal Society Open Science* **3**. (DOI:10.1098/rsos.160037).
[76] Maynard Smith, J. 1982 *Evolution and the theory of games*. Cambridge, Cambridge University Press.
[77] Iwasa, Y., Pomiankowski, A. & Nee, S. 1991 The Evolution of Costly Mate Preferences II. The 'Handicap' Principle. *Evolution* **45**, 1431-1442. (DOI:10.2307/2409890).
[78] Taylor, P. D. 1996 The Selection Differential in Quantitative Genetics and Ess Models. *Evolution* **50**, 2106-2110. (DOI:10.2307/2410769).
[79] Dieckmann, U. & Law, R. 1996 The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology* **34**, 579-612. (DOI:10.1007/bf02409751).
[80] McGill, B. J. & Brown, J. S. 2007 Evolutionary Game Theory and Adaptive Dynamics of Continuous Traits. *Annual Review of Ecology, Evolution, and Systematics* **38**, 403-435. (DOI:doi:10.1146/annurev.ecolsys.36.091704.175517).
[81] Taylor, P. D. 1981 Intra-Sex and Inter-Sex Sibling Interactions as Sex-Ratio Determinants. *Nature* **291**, 64-66. (DOI:DOI 10.1038/291064a0).
[82] Hamilton, W. D. 1967 Extraordinary sex ratios. *Science* **156**, 477-488.
[83] Lehtonen, J. & Schwanz, L. E. 2018 Mate limitation and sex ratio evolution. *Royal Society Open Science* **5**, 171135. (DOI:10.1098/rsos.171135).
[84] Lehtonen, J. & Dardare, L. 2019 Mathematical Models of Fertilization—An Eco-Evolutionary Perspective. *The Quarterly Review of Biology* **94**, 177-208. (DOI:10.1086/703633).
[85] Parker, G. A. & Pizzari, T. 2010 Sperm competition and ejaculate economics. *Biological Reviews* **85**, 897-934.
[86] Kalmus, H. 1932 Über den Erhaltungswert der phänotypischen (morphologischen) Anisogamie und die Entstehung der ersten Geschlechtsunterschiede. *Biologisches Zentralblatt* **52**, 716–736.
[87] Scudo, F. M. 1967 Adaptive value of sexual dimorphism - I, anisogamy. *Evolution* **21**, 285-291.
[88] Parker, G. A., Baker, R. R. & Smith, V. G. F. 1972 The origin and evolution of gamete dimorphism and the male-female phenomenon. *Journal of Theoretical Biology* **36**, 529-553.
[89] Wilson, D. S. & Colwell, R. K. 1981 Evolution of Sex Ratio in Structured Demes. *Evolution* **35**, 882-897. (DOI:10.2307/2407858).
[90] Nunney, L. 1985 Female-biased sex ratios: individual or group selection? *Evolution*, 349-361.
[91] Fisher, R. A. 1930 *The genetical theory of natural selection*. Oxford, Oxford University Press.

Appendix

Here we prove some identities needed in the main text. Below, *i* is an index running over all individuals in the global population, ignoring groups. *k*, in turn, indexes the groups, *j* indexes individuals within a group and *n* indicates group size. Instead of the short notation E and cov used in the main text, we will explicitly indicate group-size weighted averages and covariances following Price [5]:

$$ave_n W = \frac{\sum_k n_k W_k}{\sum_k n_k} \tag{A1}$$

$$cov_n(W, G) = \frac{\sum_k n_k (W_k - ave_n W)(G_k - ave_n G)}{\sum_k n_k} \tag{A2}$$

The aim is to prove

$$cov_n(W, G) = cov(w, G) = cov(W, g) \tag{A3}$$

and

$$ave_n[cov_k(w, g)] = cov(\Delta w, g) = cov(w, \Delta g) = cov(\Delta w, \Delta g) \tag{A4}$$

The proof for equation A3 is as follows:

$$cov_n(W, G) = \frac{\sum_k n_k (W_k - ave_n W)(G_k - ave_n G)}{\sum_k n_k} = \frac{\sum_k n_k (\frac{\sum_j w_{kj}}{n_k} - \bar{w})(G_k - \bar{g})}{\sum_k n_k} = \frac{\sum_k (\sum_j w_{kj} - n_k \bar{w})(G_k - \bar{g})}{\sum_k n_k} = \frac{\sum_k \sum_j (w_{kj} - \bar{w})(G_k - \bar{g})}{\sum_k n_k} = \frac{\sum_i (w_i - \bar{w})(G_i - \bar{g})}{N} = cov(w, G)$$

Due to symmetry, this also implies that $cov(W, g) = cov(w, G) = cov_n(W, G)$. Since this holds for weighted covariance, it naturally also holds for unweighted covariance which is a special case of the former. A further useful consequence of the above identity is $var(G) = cov(G, G) = cov(G, g)$.

Next we derive the equivalence A4:

$$ave_n\left[cov\left(w_{kj}, g_{kj}\right)\right] = ave_n\left[cov\left(w_{kj} - W_k, g_{kj} - G_k\right)\right] = \frac{\sum_k n_k \left(\frac{\sum_j (w_{kj} - W_k - 0)(g_{kj} - G_k - 0)}{n_k}\right)}{\sum_k n_k} = \frac{\sum_k \sum_j (w_{kj} - W_k - 0)(g_{kj} - G_k - 0)}{N} = \frac{\sum_i (w_i - W_i - 0)(g_i - G_i - 0)}{N} = cov(w - W, g - G) = cov(\Delta w, \Delta g) =$$

$$cov(w, g) - cov(w, G) - cov(W, g) + cov(W, G) = \qquad (**)$$

$$cov(w, g) - cov(w, G) - cov(w, G) + cov(w, G) = cov(w, g) - cov(w, G) =$$

$$cov(w, g - G) = cov(w, \Delta g)$$

Alternatively, following from (**):

$$cov(w, g) - cov(w, G) - cov(W, g) + cov(w, G) = cov(w, g) - cov(W, g) =$$

$$cov(w - W, g) = cov(\Delta w, g)$$