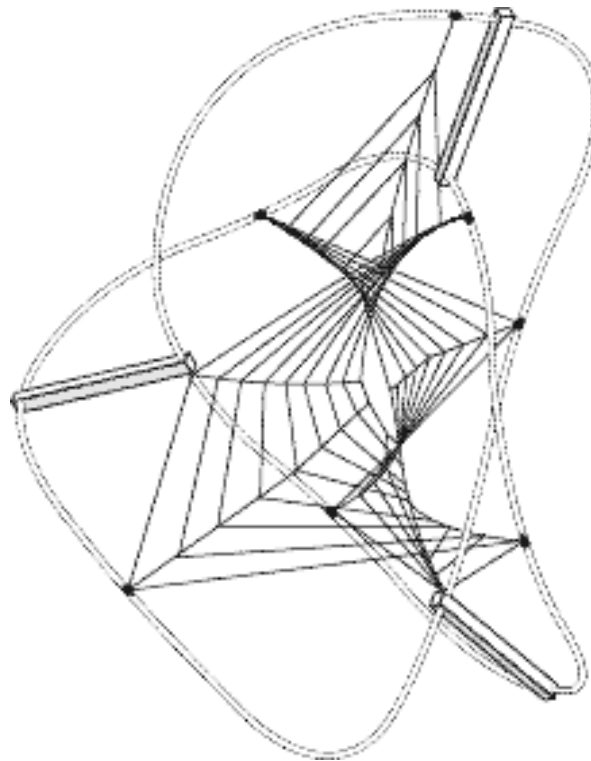


Centre for Philosophy of Natural and Social Science**Causality: Metaphysics and Methods**

Technical Report 13/03

Multi-Level Selection, Price's Equation and Causality

Samir Okasha



Editor: Julian Reiss

Multi-Level Selection, Price's Equation and Causality^{*}

Samir Okasha
Department of Philosophy
University of York
York
YORK YO10 5DD
so5@york.ac.uk

July 2003

^{*} Thanks to David Sloan Wilson, Elliott Sober, Peter Godfrey-Smith and Benjamin Kerr for detailed comments and extensive correspondence concerning these issues, to Charles Goodnight for help with contextual analysis, and to two anonymous referees for their criticisms. Thanks to the AHRB for financial support.

Abstract

Two alternative statistical approaches to modelling multi-level selection in nature, both found in the contemporary biological literature, are contrasted. The simple covariance approach partitions the total selection differential on a phenotypic character into within-group and between-group components, and identifies the change due to group selection with the latter. The contextual approach partitions the total selection differential into different components, using multivariate regression analysis. The two approaches have different implications for the question of what constitutes group selection and what does not. I argue that the contextual approach is theoretically preferable. This has important implications for a number of issues in the philosophical debate about the levels of selection.

1 Introduction

This paper deals with conceptual aspects of the ‘levels of selection’ question in evolutionary biology. I contrast two different statistical approaches to the diagnosis of group selection in nature, which I call the ‘simple covariance approach’ and the ‘contextual approach’ respectively. I examine the relation between the two approaches, focusing on a certain critical case where the two approaches disagree about whether or not group selection is occurring. I argue that the contextual approach gives the intuitively correct answer. I then consider a modification of the simple covariance approach, which comes from employing a special definition of what a ‘group’ is, due to Sober and Wilson ([1998]). I show that this modification narrows the gap between the simple covariance and contextual approaches. However, it does not eliminate it entirely: there is still a (different) critical case where the two approaches disagree about the occurrence of group selection. I argue that in this case, there are considerations in favour of both answers, though on balance, the contextual approach is theoretically preferable. This has interesting implications for the use of ‘frameshifting’ arguments in the levels of selection debate, and the much discussed issue of additivity.

2 Group Selection and the Covariance Formulation of Selection

As evolutionists have long recognized, many if not most species are sub-divided into partially isolated groups of various sorts, ranging from small family associations to larger units such as tribes, colonies, and local breeding populations or demes. The existence of this group structure can have an importance influence on the course of evolution, as numerous theoretical and empirical studies have shown, leading to outcomes which would not occur in a single panmictic population. The evolution of ‘altruistic’ or self-sarcificial behaviours is perhaps the best known example; recent work has shown the pervasive importance of population sub-division in permitting such behaviours to evolve (Sober and Wilson [1998]; Frank [1999]; Wade [1996]). Where a population is sub-divided into groups, the evolution of a trait can be influenced by selection at the group level as well as at the individual level; this is what permits evolutionary outcomes which

would be impossible in a single unstructured population, where selection can only operate on fitness differences between individual organisms.

The group selection debate has generated a vast literature, in both biology and philosophy, which I will not review here. Suffice it to say that not all authors are in agreement about whether certain processes constitute ‘real’ group selection or not, so the debate has a conceptual as well as an empirical aspect. In this and the following section, I outline two different statistical approaches to modelling group selection, employed in the contemporary biological literature. The two approaches have importantly different implications for the question of what constitutes group selection and what does not; it is these implications, rather than the technical details of the approaches, which will be the main focus here.

Consider first a single unstructured population, with no groups. For natural selection to operate, phenotypic differences between organisms must correlate with differences in their fitness. As Price ([1970]) and Robertson ([1966]) first showed, the selection differential on any phenotypic character, defined as the change in the average value of the character in the population due to selection *within* a generation, is given by:

$$\Delta \bar{z} = \text{Cov} (w_i, z_i) \quad (2.1)$$

where z_i is the phenotypic value of the i^{th} organism, w_i is the relative fitness of the i^{th} organism, and \bar{z} is the average phenotypic value in the population. Equation 2.1 tells us that the change in the average character due to selection within a generation equals the covariance between entity relative fitness and entity character value. This simply formalizes the intuitive idea that natural selection depends on an association between fitness and character - if taller organisms have a survival advantage over shorter ones, i.e. if height and fitness covary positively, then average height in the population will be expected to increase.

Three points about equation 2.1 should be noted. First, the equation requires that the phenotypic character z be assigned a metric, so it applies in the first instance to ‘continuous’ characters such as height, rather than ‘discrete’ characters such as eye-colour. However, this is not a serious limitation. Suppose eye-colour is the trait

undergoing selection, and we are interested in the proportion of brown-eyed organisms in the population. We simply define $z_i = 1$ if the i^{th} organism is brown-eyed, and $z_i = 0$ otherwise. So the value of z for an organism indicates whether it is brown-eyed or not. Obviously, \bar{z} then equals the proportion of brown-eyed organisms in the population. So equation 2.1 can be applied: it tells us that the change in the proportion of brown-eyed organisms due to selection within a generation equals the covariance between fitness and z , where z is defined as above. Similarly, z could be defined as the frequency of a particular allele at a given locus in an organism ($= 1, \frac{1}{2}$ or 0 for diploid organisms); \bar{z} would then equal the overall frequency of the allele in the population. Equation 2.1 can therefore handle discrete as well as continuous traits.

Secondly, note that equation 2.1, which is sometimes called the ‘basic Price equation’, concerns only selection itself, not the response to selection; the latter depends on the fidelity of transmission across generations, hence on facts about the inheritance mechanism etc.¹ Throughout this paper, our concern will be with selection itself, hence the within-generation change in \bar{z} , rather than the response to selection. So we do not need to worry about the genetic basis of z , whether it is heritable, and so on. This separation of selection from the response to selection is standard fare in quantitative genetics.

Thirdly, note that equation 2.1 is statistical not causal. A positive value of $\text{Cov}(w_i, z_i)$ tells us that organisms with a high value of z are favoured by selection, but it does not imply a direct causal link between fitness and z . It is possible that z itself has no causal effect on fitness, but is closely correlated with another phenotypic character which does causally influence fitness. Despite this limitation, equation 2.1 is a useful conceptual resource for students of evolution. For in general, we cannot directly test causal models of the processes affecting natural populations; we can only directly test the statistical relationships that our causal models imply. A positive value of $\text{Cov}(w_i, z_i)$ is evidence that z causally affects fitness, but such evidence is of course defeasible.

Now we introduce population structure. Suppose our population is divided into a number of groups, as depicted in the Figure below. For simplicity, we assume that each

¹ The full Price equation, of which the equation above is a simplified version, does deal with transmission across generations. See Frank ([1999]) for a good introduction.

group contains the same number of organisms. Nothing in particular about the nature of the groups is assumed – they could be kin groups, or colonies, or demes, or anything else. It may help to think of the groups as geographically separate, but this is only a heuristic device for understanding the formalism below; it is not required by the formalism itself. (Indeed, the formalism would work perfectly well even if the groups were simply arbitrarily defined by the investigator, and had no biological reality at all.) The nature of the groups, and the question of their biological justification, is an issue we return to below.

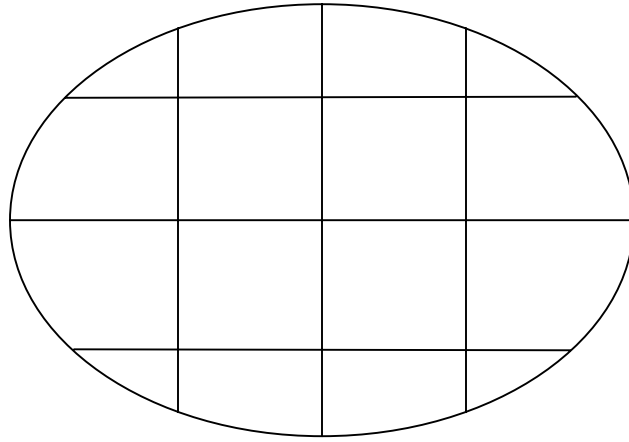


Figure 1: A Sub-divided Population

We define z_{ij} as the phenotypic value of the i^{th} organism in the j^{th} group, and w_{ij} its relative fitness (relative to all the other organisms in the whole population). We define Z_j as the average phenotypic value of the organisms in the j^{th} group, and W_j as the average relative fitness of the organisms in the j^{th} group. So $Z_j = \frac{1}{n} \sum_i z_{ij}$ and $W_j = \frac{1}{n} \sum_i w_{ij}$, where n is group size. We can think of Z_j as the ‘group phenotype’ of the j^{th} group, and W_j as its ‘group fitness’.

Since the population is sub-divided into groups, it is possible to partition the total selection differential on the character z , given in equation 2.1 above, into within-group and between-group components, as Price [(1972)] showed. This gives:

$$\Delta \bar{z} = E[\text{Cov}_{\mathbf{W}}(z_{ij}, w_{ij})] + \text{Cov}_{\mathbf{B}}(Z_j, W_j) \quad (2.2)$$

The first term on the RHS of equation 2.2, $E[\text{Cov}_{\mathbf{W}}(z_{ij}, w_{ij})]$, is the average of the *within-group* covariances between z and w . To calculate it, we consider each group separately, and calculate the covariance between fitness and character value within that group; we then take the average across all the groups. The second term on the RHS, $\text{Cov}_{\mathbf{B}}(Z_j, W_j)$, is the covariance between group phenotype Z_j and group fitness W_j ; to calculate it, we simply calculate the value of W_j and Z_j for each of the groups, then compute their covariance.² Equation 2.2 involves no new biological assumptions; it follows directly from equation 2.1, given that the population is sub-divided into groups. (See Wade ([1985]) or Price ([1972]) for a full derivation of equation 2.2.)

Numerous authors have regarded equation 2.2 as a useful way of modelling the combined effects of group selection and individual selection on a character. The equation appears to partition the total selection differential on a character into two components, corresponding respectively to individual selection and group selection. Price ([1972]), Hamilton ([1975]), Wade ([1985]) and others have all endorsed this suggestion.³ On their view, $E[\text{Cov}_{\mathbf{W}}(z_{ij}, w_{ij})]$ represents the effect of individual selection on the change in \bar{z} , while $\text{Cov}_{\mathbf{B}}(Z_j, W_j)$ represents the effect of group selection. I call this the ‘simple covariance approach’ to group selection.

The grounds for the simple covariance approach are easy to see. Suppose that all the groups have the same fitness, or that group fitness and group character are uncorrelated. Intuitively there can be no group selection in such a situation. And the term $\text{Cov}_{\mathbf{B}}(Z_j, W_j)$ will equal zero, in that situation. Conversely, suppose that within each group, all organisms have the same fitness, or that organismic fitness and character are

² The term $\text{Cov}_{\mathbf{B}}(Z_j, W_j)$ would need to be weighted by the number of individuals in each group, if we had not made the simplifying assumption that all groups are equal in size. Therefore, $\text{Cov}_{\mathbf{B}}(Z_j, W_j)$ is really an expectation taken over all the individuals, not the groups.

³ Thought in more recent writings, Wade appears sympathetic to the alternative contextual approach described below; see Wade, Goodnight and Stevens ([1999]).

uncorrelated within each group. Intuitively there can be no individual selection in such a situation. And the term $E[\text{Cov}_w(z_{ij}, w_{ij})]$ will be zero in that situation – for each of the within-group covariances will be zero, so their average will be zero.⁴ In other words, given the intuitively plausible idea that group selection is selection *between* groups, and individual selection is selection *between* individuals within the *same* group, equation 2.2 tells us how much of the change in \bar{z} is due to group and individual selection respectively. It is easy to see why equation 2.2 is widely regarded as a valuable tool for modelling multi-level selection.

3 The Contextual Approach

The second approach to the diagnosis of group selection, which I call the contextual approach, is motivated by an apparent shortcoming of the simple covariance approach which Heisler and Damuth ([1987]) and Nunney ([1985]) have emphasized; Sober ([1984]) made a closely related point in a slightly different context. The shortcoming can be seen as follows. Consider again a population sub-divided into groups of equal size. Suppose that the fitness of a given organism depends only on its own phenotype, and not on the composition of its group. So any two organisms of identical phenotype have the same fitness, irrespective of the groups in which they are found – there are no ‘group effects’ on individual fitness. As Heisler and Damuth ([1987]) note, most biologists would say that no group selection is occurring in such a situation, for the evolution of the system can be predicted *without* taking group structure into account. Nonetheless, the term $\text{Cov}_B(Z_j, W_j)$ in equation 2.2 may very well be non-zero. Some groups may be fitter than others, simply because they contain a higher proportion of fitter organisms. So equation 2.2 detects group selection where intuitively there is none.

Sober ([1984]) illustrated this problem with an example in which an organism’s fitness depends positively on its own height, but is independent of the composition of its group - any two organisms with the same height have identical fitness, whatever group

⁴ However, in general one cannot infer that if $E[\text{Cov}_w(z_{ij}, w_{ij})] = 0$, there is no selection at the individual level. The most one can infer (given the suggestion in question for how to determine the levels of selection) is that individual level selection has no *net* effect on the frequency of the A particle. This is compatible with the existence of individual level selection, if the direction of the selection is different in different groups.

they are found in. So there are no ‘group effects’ on individual fitness – fitness is a function of individual phenotype alone. Nonetheless, there are fitness differences between groups, because the groups vary with respect to the average height of their constituent organisms. A group composed of tall organisms will be fitter than a group composed of short organisms. Sober concluded from this example that between-group variance in group fitness (average individual fitness) is not a sufficient condition for group selection. Though Sober’s discussion was not framed in terms of the Price equation, his example creates an immediate problem for the simple covariance approach to group selection. If we agree that there is no group selection in Sober’s example, as most people would, then equation 2.2 cannot be taken to partition the total selection differential into components corresponding to individual and group selection respectively.

As a solution to this problem, Heisler and Damuth ([1987]) advocate a statistical technique drawn from the social sciences called ‘contextual analysis’.⁵ The basic idea is to regard group character as a ‘contextual’ or relational character of each individual in the group. So in the example above, the average height of a group is regarded as a contextual character of each individual in that group. Each individual in the population is therefore assigned two characters: an individual character (height), and a contextual character (average height of its group). Both of these characters are capable of affecting an organism’s fitness. The crucial question vis-à-vis group selection is: does the contextual character (group character) of an organism affect its fitness? Only if the answer is ‘yes’, should we say that group selection is in operation – for only then is there a group-level effect on fitness.

How should we determine the answer to this question? Simply looking for a correlation between fitness and group character is insufficient, as Heisler and Damuth stress. Even if an organism’s group character does *not* affect its fitness, there will still be a correlation between fitness and group character *so long as there is a correlation between fitness and individual character*. This is because individual character and group character are *themselves* correlated – taller organisms are more likely to be found in

⁵ See Boyd and Iversen ([1979]) for a thorough introduction to contextual analysis; see Heisler and Damuth ([1987]), Goodnight et. al. ([1997]), and Tsuji ([1995]) for applications to levels of selection problems in biology.

groups with high average height, obviously.⁶ To determine whether group character affects individual fitness, we need to determine whether there is a correlation between fitness and group character that is not due to the correlation between fitness and individual character. In other words, does the group character of an individual help predict its fitness, *over and above* the prediction that can be made on the basis of the individual's own phenotypic character? To answer this question, we employ the following linear regression model:

$$w_{ij} = \beta_1 z_{ij} + \beta_2 Z_j + e_{ij}$$

where β_1 is the partial regression of individual fitness on individual character, controlling for group character, β_2 is the partial regression of individual fitness on group character, controlling for individual character, and e_{ij} is the residual whose variance is to be minimized. Therefore, β_2 is the change in individual fitness that would result if the group character of an individual *of fixed phenotype* were changed by one unit – it measures the extent to which differences in group character predict differences in individual fitness, holding individual character constant.

Heisler and Damuth's suggestion is that group selection requires that β_2 be non-zero. This means that information about the group to which an individual belongs is relevant to predicting the individual's fitness, over and above information about the individual's own phenotype, i.e. it signals a 'group effect' on individual fitness. In Sober's ([1984]) example above, where the fitness of an individual depends only on its own height, then β_2 is zero – once you know the height of an individual organism, further information about its group character does not help you make a better prediction about its fitness. Of course, if you did *not* know the individual's height then being told its group character *would* help you predict its fitness, so the *simple* regression of fitness on group character is non-zero. But group character is not a significant predictor of individual fitness, once individual character has been taken into account. That is the crucial point.

⁶ Individual character and group character will always be correlated if group character is defined as average individual character, except in the limiting case where all groups have the same group character. However, the method of contextual analysis is applicable to cases where group character is not defined this way, in which case the existence of a correlation between individual and group character is an empirical matter.

So the contextual analysis approach yields the intuitively correct result – that where individual fitness depends only on individual phenotype and not on group composition, there is no group selection. In this respect, it seems clearly superior to the simple covariance approach.

As Heisler and Damuth ([1987]) note, the contextual analysis approach corresponds quite closely to Sober’s ([1984]) probabilistic treatment of group selection. Sober ([1984]) argued that for there to be group selection, it must be true that group character is a ‘positive causal factor’ in determining organismic fitness. According to Sober, this means that if an organism of given individual character was moved to a group with a different group character, its probability of survival/reproduction (i.e. its fitness) would be altered. In the contextual analysis framework, this implies that the partial regression of fitness on group character should be non-zero.⁷ So although Sober’s ([1984]) approach, unlike contextual analysis, is formulated in causal rather than statistical terms, and although Sober does not try to *quantify* the amount of the total selection differential that is due to group selection, while contextual analysis does, the underlying idea is similar in both cases.⁸

It helps to explore the relationship between the simple covariance and the contextual approaches in exact terms. As we saw, the simple covariance approach partitions the total selection differential into two components:

$$\Delta \bar{z} = E[\text{Cov}_{\mathbf{W}}(z_{ij}, w_{ij})] + \text{Cov}_{\mathbf{B}}(Z_j, W_j) \quad (2.2)$$

The contextual approach also partitions the total selection differential, but into different components:

$$\Delta \bar{z} = \beta_1 \text{Var}(z_{ij}) + \beta_2 \text{Var}(Z_j) \quad (2.3)$$

The first term on the RHS of equation 2.3, $\beta_1 \text{Var}(z_{ij})$, is the partial regression of individual fitness on individual character, times the variance in individual character; the second term, $\beta_2 \text{Var}(Z_j)$, is the partial regression of individual fitness on group character,

⁷ Though the converse is not the case. If the partial regression of fitness on group character is non-zero, this does not necessarily make group character a ‘positive causal factor’ in Sober’s ([1984]) sense.

⁸ Thanks to Elliott Sober (personal communication) for pointing out this similarity.

times the variance in group character. These terms correspond respectively to the change in \bar{z} due to individual and group selection, according to the contextual approach. Obviously, $\beta_2 \text{Var}(Z_j)$ will only be non-zero if both $\text{Var}(Z_j)$ and β_2 are non-zero; so group selection requires $\beta_2 \neq 0$. (See Heisler and Damuth ([1987]) or Goodnight et. al. ([1997]) for a derivation of equation 2.3).

The relation between equations 2.2 and 2.3 can be seen more clearly by rearranging equation 2.2. Note firstly that since group fitness W_j equals mean individual fitness, the term $\text{Cov}_B(Z_j, W_j)$ equals $\text{Cov}(Z_j, w_{ij})$, where the latter is the covariance between individual fitness and group character of the individual. Secondly, in virtue of the general relation between covariance and regression, $\text{Cov}(Z_j, w_{ij}) = b_{wZ} \text{Var}(Z_j)$, where b_{wZ} is the *simple* regression of individual fitness on group character. Therefore, equation 2.2 can be written:

$$\Delta \bar{z} = E[\text{Cov}_w(z_{ij}, w_{ij})] + b_{wZ} \text{Var}(Z_j)$$

This permits simple comparison between the two approaches with respect to the criterion for group selection. On the simple covariance approach, the condition for group selection is that $b_{wZ} \text{Var}(Z_j) \neq 0$. This in turn requires that $b_{wZ} \neq 0$. On the contextual approach, group selection requires $\beta_2 \neq 0$. Therefore, the simple covariance approach says that group selection requires the *simple* regression of individual fitness on group character to be non-zero, while the contextual approach says that the *partial* regression of individual fitness on group character must be non-zero. In the example above, where an organism's fitness depends only on its own phenotype, the *simple* regression of fitness on group character is non-zero, for differences in group character do help predict differences in individual fitness. However, the *partial* regression of fitness on group character is zero – for the association between individual fitness and group character goes away, once we control for individual character. So the contextual approach detects no group selection, while the simple covariance approach does. Intuitively, the former is surely the correct answer.

The conditions $b_{wZ} \neq 0$ (simple covariance approach), and $\beta_2 \neq 0$ (contextual approach) should be interpreted as alternative *necessary* conditions for group selection, not necessary and sufficient conditions. Group selection is meant to be a causal process in

nature, and as noted above, no purely statistical approach can define, or provide necessary and sufficient conditions for, the occurrence of a causal process. But taken as just necessary conditions, $\beta_2 \neq 0$ and $b_{wZ} \neq 0$ correspond to alternative conceptions of what the causal process of group selection amounts to. If we take group selection to require $b_{wZ} \neq 0$, this means that we take group selection to be a process in which groups with different group characters proliferate differentially, whether or not individual fitness is causally affected by group character.⁹ If we take group selection to require $\beta_2 \neq 0$, this means that we take group selection to be a process in which fitness differences between organism are caused, in part, by differences in group character. In neither case do the statistical inequalities provide sufficient conditions for group selection.

Of course, a defender of the simple covariance approach might simply reject the intuition that where individual fitness depends only on individual phenotype, there is no group selection. But this move is not especially plausible. As the history of the group selection debate shows, the existence of group-level effects on individual fitness has always been at the heart of the issue. Much of the interest in group selection has stemmed from its potential to explain altruism in nature; and where altruistic traits are at issue, there are by definition group-level effects on individual fitness – an individual’s fitness is boosted by the presence of other altruists in its group. Most evolutionists would agree that group selection can also operate on non-altruistic traits, i.e. individual and group selection *need* not oppose each other, but the centrality of altruism to the group selection debate cannot be denied. So an approach which permits group selection to operate in the absence of group-level effects on individual fitness, such as the simple covariance approach, seems clearly at odds with the traditional explanatory role of the concept.

4 A Modification of The Simple Covariance Approach

In their recent book, Sober and Wilson ([1998]) endorse the covariance approach – they explicitly allude to equation 2.2 as a way of separating the effects of group and individual

⁹ Importantly, ‘proliferate differentially’ here means contribute differentially to the next generation of *organisms*, not groups. This is because group fitness is defined as average individual fitness, not expected number of offspring groups. See the discussion in section 4 below concerning the distinction between group selection type 1 (GS1) and group selection type 2 (GS2).

selection.¹⁰ But unlike other supporters of simple covariance, Sober and Wilson offer a very careful definition of what a ‘group’ is. They do not identify groups geographically, but rather on the basis of fitness interactions. An organism’s group consists of those other organisms with which it has fitness-affecting interactions. This implies that groups must be defined on a trait-by-trait basis, they argue – hence the concept of a ‘trait group’ (first introduced by Wilson [1975]). Suppose body size is the trait of interest. An organism’s trait group then consists of all those other organisms whose body size affects its fitness. The organism’s trait group for size will very probably be different from its trait group for eye-colour, or for parental care, for example. Sober and Wilson argue at length that the trait group concept is the appropriate one, for the group selection issue. There is a biological rationale for dividing up a population into trait groups, that there is not for geographical groups, or groups individuated in any other way. Indeed, Wilson (personal communication) holds that the trait group concept has in fact been implicit in many biologists’ discussions of group selection all along.

Given the Sober and Wilson definition of a group, the difference between the simple covariance approach and the contextual approach is considerably narrowed. To see this, consider again the case where an organism’s fitness depends on its own phenotype alone, irrespective of group composition – the case that proved embarrassing for the simple covariance approach above. In this situation, Sober and Wilson will say that *no groups in fact exist* with respect to the phenotypic trait in question. This obviously implies that there is no group selection – which is exactly what the contextual approach says. So if advocates of the simple covariance approach employ the Sober and Wilson trait group concept, their approach becomes significantly closer to the contextual approach. Indeed, the technique of contextual analysis can be thought of as a way of diagnosing the existence of groups in a population, in the Sober and Wilson sense of ‘group’.

¹⁰ There is a slight exegetical complication here. For the most part, Sober and Wilson’s discussion is framed in terms of the covariance approach. However, in a footnote on p. 343 of *Unto Others*, they say that the contextual approach is actually superior, but that the difference between contextual and covariance approaches is unimportant for their purposes. However, Sober and Wilson (personal communication) both insist that in the ‘soft selection’ model discussed below, there is no group selection; since the contextual approach says that there *is* group selection in this model, while the covariance approach denies it, this justifies my attribution of the latter position to Sober and Wilson, despite the footnote in question.

Where the simple covariance approach to group selection is employed in conjunction with the Sober and Wilson definition of a ‘group’, I call the result the ‘modified covariance approach’ to group selection. The modified covariance approach agrees with the contextual approach that where individual fitness is a function of individual phenotype alone, there is no group selection, i.e. group-level effects on fitness are a necessary condition of group selection. Does this mean that the modified covariance and the contextual approaches are fully equivalent? It turns out that the answer is no.

Their non-equivalence can be seen by considering a type of selection known as ‘soft selection’ (Wade [1985], Goodnight et. al. [1997]). As before, a population is subdivided into groups of equal size. We assume that the ‘groups’ in question are *bona fide* trait-groups, identified according to the Sober and Wilson criterion of fitness interaction. In soft selection, all of the groups have identical fitness. This could occur if resource limitations constrain each group to leave exactly the same number of progeny. Nonetheless, there are group effects on individual fitness. The fitness of any individual depends not just on its own phenotype, but also on its phenotypic ranking within its group. Suppose the phenotypic trait is ‘body size’. An organism’s fitness then depends not just on its own body size, but also on its size relative to the other members of its group. Larger organisms have a fitness advantage relative to smaller ones within any group, e.g. because they can appropriate more of the group’s resources for themselves. So any organism, large or small, benefits from being in a group of low average size. So if a given organism of fixed size is moved from a group with high average body size to a group with low average body size, its fitness will increase. Fitnesses within each group are therefore strongly frequency-dependent.

In this example, the modified covariance approach implies that there is no group selection. By design, all the groups have equal fitness, hence the covariance between group fitness and group character automatically equals zero, so $b_{wZ} = 0$. However, the contextual approach detects a component of group selection, as Goodnight et. al. ([1997]) have shown. In soft selection, β_2 is non-zero – information about group character does predict individual fitness, controlling for individual phenotype. This is easy to see intuitively. Since two organisms with the same individual character will differ in fitness if their group characters are different, differences in group character will help predict

differences in individual fitness, controlling for individual character. Hence β_2 will be non-zero. So the modified covariance approach and the contextual approach yield different verdicts about whether group selection is occurring or not. The former says ‘no’, the latter ‘yes’.

These conflicting verdicts can be illustrated with a simple numerical example of soft selection. Imagine a population of fifty organisms, divided into five trait groups of size ten. Organisms are of two types, large (L) or small (S). We define z , the individual character, as follows: $z = 1$ if an organism is small, 0 otherwise. So \bar{z} is the overall frequency of small organisms in the population. The initial value of \bar{z} is $\frac{1}{2}$. We define Z , the group character, as the average value of z within a group. The composition of the five groups is shown in the Table below. Each group has an *absolute* fitness of 20, i.e. contributes 20 organisms to the next generation. But this group output is shared unequally between large and small organisms, the former getting more of it. So within any group, large organisms are fitter. Additionally, there are group effects on individual fitness – any organism, whether large or small, benefits from being in a group with a larger value of Z , i.e. one containing more small organisms. This is shown in the Table below, where w_L and w_S denote the within-group *absolute* fitnesses of large and small organisms respectively. In groups 1 and 2, which have very low values of Z , a small organism has a fitness of zero – it does not share in the group output. But in group 5, which has a high value of Z , small organisms have a fitness of $14/9$. Similarly, the fitness of a large organism increases as we move from group 1 through group 5.

	Group Composition	Group Output	Group Character(Z)	w _L	w _S
Group 1	9L, 1S	20L, 0S	0.1	20/9	0
Group 2	8L, 2S	20L, 0S	0.2	20/8	0
Group 3	5L, 5S	15L, 5S	0.5	3	1
Group 4	2L, 8S	10S, 10A	0.8	5	1.25
Group 5	1L, 9S	6S, 14A	0.9	6	14/9

Table 1: Soft Selection

We are interested in $\Delta \bar{z}$. The basic Price equation 2.1 tells us that $\Delta \bar{z} = \text{Cov}(w_i, z_i)$, where w_i is the *relative* fitness of the i^{th} organism. Applied to the data in Table 1, this gives:

$$\Delta \bar{z} = \text{Cov}(w_i, z_i) = -0.21$$

So the overall selection differential on the character z is -0.21 , i.e. within-generation selection will reduce the frequency of small organisms from 0.50 to 0.29. Now consider the covariance approach, which partitions the selection differential according to:

$$\Delta \bar{z} = E[\text{Cov}_{\mathbf{W}}(z_{ij}, w_{ij})] + \text{Cov}_{\mathbf{B}}(Z_j, W_j)$$

Applied to the data in Table 1, this yields:

$$\Delta \bar{z} = -0.21 + 0 = -0.21$$

So on the covariance approach, there is no group selection, just as expected, since all the groups have the same fitness.

Now consider the contextual approach:

$$\Delta \bar{z} = \beta_1 \text{Var}(z_{ij}) + \beta_2 \text{Var}(Z_j)$$

Applied to the data in Table 1, this yields

$$\Delta \bar{z} = -0.35 + 0.14 = -0.21$$

So contextual analysis detects a component of both group selection and individual selection. Individual selection reduces the value of \bar{z} while group selection raises it; the former is stronger, so \bar{z} declines overall. This is just what we should expect – a high value of z reduces an individual's fitness, but being in a group with high mean z (i.e. high Z) increases individual fitness. In soft selection, the contextual and covariance approaches give different answers to the question 'is group selection occurring or not?'.

Which answer is correct? This is a tricky question. A case can be made in favour of both answers. If we accept the contextual approach, this commits us to saying that group selection can occur even if there is no variance in group fitness. Intuitively this may sound strange – surely selection at any given hierarchical level requires variance in fitness at that level, as Lewontin ([1970]) famously argued? This is a point in favour of the modified covariance approach. On the other hand, the contextual approach seems superior on theoretical grounds. If we think of individual selection and group selection as separate evolutionary 'forces', both capable of influencing the evolution of a character, the use of partial regression techniques is surely appropriate. The basic motivation behind the contextual approach – that an association between fitness and group character may be a mere side effect of the association between fitness and individual character – is hard to deny; it is simply a special case of the familiar point that direct selection on a given character will have indirect effects on any correlated characters (Pearson [1903], Lande and Arnold [1983]). Considering the partial regression of fitness on a character, controlling for correlated characters, rather than the simple regression, is the obvious way to accommodate this point. This favours the contextual approach.

My own view is that the contextual approach is superior, despite the apparently anomalous result that group selection can occur without variance in group fitness. (More on this anomaly in a moment.) Importantly, this in no way mitigates against the Sober and Wilson definition of a 'group'. It is quite possible to accept that the trait group concept provides the appropriate way of dividing up a population into groups, while insisting that the partial regression of individual fitness on group character, not the simple regression, is the relevant quantity for determining whether group selection is operating.

(This point applies however we define groups.) Whether one favours the covariance approach or the contextual approach, some way of dividing a population up into groups is necessary. Equations 2.2 and 2.3 do not in themselves tell us how this division should be performed; rather they presume that it has already been done. So advocates of the contextual approach need some criterion for identifying the groups in a population, no less than advocates of the covariance approach. The trait group concept is an obvious candidate, for it has a clear biological justification. In short, we can accept Sober and Wilson's analysis of what a group is without accepting their criterion for when group selection is occurring.

In fact, a case can be made that anyone who accepts Sober and Wilson's trait group concept *should* favour the contextual, rather than the covariance, approach. The rationale behind the trait group concept is that where there are fitness-affecting interactions between organisms, the organisms so affected may together constitute an evolutionary significant unit. This suggests that the identification of group-level effects on individual fitness is the critical issue – we are interested in situations where an organism's fitness depends on its interactions with other organisms, not just on its own phenotypic traits. But if that is so, then it is natural to suggest that group selection is occurring whenever individual fitness is *directly* affected by group membership, i.e. when there is an association between fitness and group character than is not just a side-effect of the association between fitness and individual character – just as the contextual approach does. The main argument in favour of the trait group criterion for what a group is, is also an argument in favour of the contextual criterion for when group selection is occurring. To put the point another way, if one uses the covariance rather than the contextual criterion for group selection, then there is no theoretical rationale for defining groups on the basis of fitness-interactions, rather than geographically, or in some other way.

What about the apparent anomaly – that the contextual approach implies that group selection can occur without variance in group fitness? This anomaly *sounds* serious – for the Lewontin principle that selection at a given level requires variance in fitness at that level is virtually axiomatic – but in fact it can be easily explained. The key point is that there are two different types of group selection. In group selection type 1 (GS1), the fitness of a group is defined as the average (or total) fitness of its constituent organisms,

i.e. the group's expected contribution to the next generation of organisms. In group selection type 2 (GS2), the fitness of a group is defined as its expected number of offspring groups, i.e. its expected contribution to the next generation of groups. This distinction has been made by a number of authors (e.g. Damuth and Heisler [1988]; Arnold and Fristrup [1982]; Sober [1984]; Okasha [2001], [2003] and others.) The GS1 and GS2 definitions of group fitness are not equivalent; which one is appropriate depends on what we want to explain. A GS2 hypothesis explains the changing frequency of different types of *group* in a metapopulation of groups, while a GS1 hypothesis explains the changing frequency of different types of *organism* in a population that is subdivided into groups. The two types of group selection thus have fundamentally different *explananda*.

The type of group selection we have been dealing with above is obviously GS1: we explicitly defined group fitness W_j as average individual fitness. (As Damuth and Heisler ([1988]) note, most models of group selection are of the GS1 rather than the GS2 type.) Both the contextual approach and the simple covariance approach are restricted to GS1; neither can be applied to GS2. The distinction between GS1 and GS2 helps explain the apparent anomaly above. When it is said that selection at a given hierarchical level requires variance in fitness at that level, this is an ambiguous assertion. Is the fitness of a given higher-level entity defined in the GS2 way, as its expected number of offspring higher-level entities, or is it defined in the GS1 way, as the average fitness of the lower-level entities that it contains? If the former, then variance in fitness at a level is certainly necessary for selection at that level. If the latter, then it is not *obvious* that selection requires variance in fitness – this is precisely the point at issue between the covariance and the contextual approaches to group selection. So if we accept the contextual approach, this does not force us to abandon the revered Lewontin principle altogether. Rather, it forces us to restrict that principle to cases where an entity's fitness is defined in the GS2 way, rather than the GS1 way.

The foregoing arguments do not provide *conclusive* grounds for favouring the contextual over the modified covariance approach.¹¹ Indeed, a knockdown argument

¹¹ See Okasha [forthcoming] for further discussion of the differences between, and respective merits of, the contextual and covariance approaches.

either way is impossible, for it is always open to a defender of either approach to simply stipulate how he wishes to apply the term ‘group selection’. But the issue is not merely terminological; there can be better and worse reasons for using one set of concepts rather than another to describe the world. Furthermore, as I show in the section below, the choice between the contextual and the covariance approaches has a number of interesting ramifications for the levels of selection debate.

5 Consequences: Frameshifting and Additivity

5.1 Frameshifting

A pattern of argument employed repeatedly in the levels of selection debate appeals to what may be called ‘frameshifting’ – applying criteria drawn from one levels of selection question (e.g. group versus individual selection) to an analogous question at a different level (e.g. individual versus genic selection). Thus for example, Walton ([1990]) criticized Sober ([1984]) on the grounds that Sober’s criteria for determining the levels of selection did not permit frameshifting between levels.¹² In my view it is not *obviously* true that frameshifting is a *sine qua non* of any general approach to the levels of selection; there may well be asymmetries in the biological hierarchy which demand that different criteria be applied at different levels. Here is not the place to attempt a general analysis of the validity of frameshifting arguments. My interest lies specifically in frameshifting in relation to the dispute between the covariance and contextual approaches to group selection.

As a number of authors have noted, there is a formal analogy between diploid population genetics and (GS1 type) group selection (Wilson [1990]; Sober and Wilson [1998]; Kerr and Godfrey-Smith [2002]; Okasha [2004]). If we think of the diploid genotypes as ‘groups’ each containing two alleles, the formal apparatus of group selection theory applies neatly – genotypes corresponds to groups, and alleles corresponding to individual organisms. In this way, the group versus individual selection

¹² Sober and Wilson ([1994]) accepted Walton’s criticism of Sober’s earlier work, agreeing with Walton that frameshifting is indeed a condition of adequacy of any general approach to the levels of selection.

issue becomes analogous to the organismic (genotypic) versus genic selection issue.¹³ So there is an opportunity for frameshifting. Suppose firstly that we adopt the covariance approach to group selection of equation 2.2. Frameshifted downwards, this means that genotypic selection occurs when there is variation in fitness *between genotypes*, and genic selection occurs when there is variation in fitness between genes *within* genotypes. So in other words, genic selection only occurs when there is segregation distortion in heterozygotes; if segregation is Mendelian, then there can be no genic selection – all the selection must be at the genotypic level.

This conclusion, which is explicitly endorsed by Sober and Wilson ([1998]), is quite plausible. Though some ‘genic selectionists’ have been inclined to say that *all* selection is genic selection (e.g. Dawkins [1976]), it is now widely accepted that this way of putting things is unhelpful at best, reflecting a product/process confusion.¹⁴ As many authors now use the term, ‘genic selection’ refers to selection between genes within the same organism, as in cases of segregation distortion, while organismic or genotypic selection refers to selection between organisms. (This is precisely how Maynard Smith and Szathmary ([1995]) use the term ‘genic selection’ in their discussion of intra-genomic conflict.) So frameshifting the covariance approach downwards yields intuitively plausible results. It implies that genic selection requires fitness differences between genes within individual organisms, which corresponds precisely to one standard conception of what ‘genic selection’ means.

However, what if we favour the contextual approach to group selection? Frameshifted downwards, it produces a most implausible resolution of the genic versus genotypic selection debate. Consider firstly a situation analogous to the ‘soft selection’ model above. Suppose there are three diploid genotypes, AA, AB and BB, with identical fitnesses. But suppose that there are ‘genotypic effects’ on genic fitness – an A allele in an AA homozygote has lower fitness than an A allele in an AB heterozygote. This means that segregation in the heterozygote is being distorted in favour of the A allele. Most

¹³ I assume here that the questions ‘is selection at the genic or the genotypic level?’ and ‘is selection at the genic or organismic level?’ are equivalent. See Falk and Sarkar ([1992]) for an interesting argument against treating these questions as equivalent.

¹⁴ All or most selection processes result in gene frequency changes; but this does not mean that the selection process itself occurs at the genic level. Hull’s famous replicator/interactor distinction was precisely designed to capture this point.

people would say that all the selection is at the genic level, in this example – for the genotypes themselves do not differ in fitness. However, the contextual approach will say that there is a component of genotypic fitness, for differences in genotypic character will help predict differences in genic fitness, controlling for genic character. So β_2 in equation 2.3 will be zero. Just as in the soft selection model, there is selection at the higher-level despite the higher-level entities not varying in fitness.

This is not the only unpalatable consequence of the contextual approach when frameshifted downwards. Consider the following hypothetical example. Genotypic absolute fitnesses are $w_{AA} = 16$, $w_{AB} = 12$, $w_{BB} = 8$. Segregation is distorted in the heterozygote in favour of the A allele in the ratio of 8 : 4, i.e. of the 12 gametes that an AB organism is expected to contribute to the next generation, 8 are A and 4 are B. Given this fitness scheme, the contextual approach implies that all the selection is at the genic level. For the fitness of a gene is independent of its genotypic context – an A gene has an absolute fitness of 8, irrespective of the genotype it is found in, and a B gene has a fitness of 4, irrespective of the genotype it is found in.¹⁵ This is analogous to the case discussed above, where the fitness of an organism is a function of its own phenotype alone, independent of the group it is in. As we saw, in such a case, the contextual approach implies that all the selection is lower-level. So if we frameshift downwards, we are forced to conclude that genic selection is the only force in operation, given the postulated genotypic fitnesses and pattern of segregation distortion. Again, this is intuitively implausible.

In short, the covariance approach appears to frameshift down quite well, the contextual approach very badly. If one were convinced that frameshifting is a general desideratum of an approach to the levels of selection, one might regard this as a point in favour of the covariance approach to group selection. But equally, one could conclude that the analogy between group selection theory and diploid population genetics is not a very strong one. It is true that there is a formal isomorphism between them, but formal isomorphisms can obscure biologically important distinctions (cf. R. Wilson [forthcoming]). In the group selection case, the question we are critically interested in is

¹⁵ The crucial feature of this example is that segregation in the AB heterozygote is distorted in favour of the A allele in the ratio w_{AA}/w_{BB} . Wherever this condition is satisfied, then genic fitness will be independent of genotypic context.

whether there are group-effects on individual fitness. In the diploid population genetics case, we are not especially interested in the question of whether there are genotypic-effects on genic fitness. The situation described in the paragraph above, where $w_{AB} = 16$, $w_{AB} = 12$, $w_{BB} = 8$ and segregation is distorted in the ratio 8 : 4 in favour of A, is of no theoretical significance at all. This is because the explanation of *why* the fitness of an A gene is the same, whatever its genotypic context, involves two quite disparate circumstances: the fact that segregation is distorted in a certain very specific way, and the fact that genotypic fitnesses are as they are. By contrast, where the fitness of an organism depends on its own phenotype alone, irrespective of group character, this *is* theoretically significant – it signals the absence of group effects on fitness. The formal analogy between diploid population genetics and group selection theory should not be overdone.

A somewhat different way to put the point is this. The formal analogy masks the following disanalogy. In group selection theory, fitnesses are possessed in the first instance by individual organisms; ‘group fitness’, defined as average or total individual fitness, is derivative – a group only has a given fitness value in virtue of the fitnesses of its constituent organisms. In diploid population genetics it is the other way round. It is the individual organisms – the higher-level entities – which possess fitnesses in the first instance. The fitness of a gene within an organism, defined as the number of that organism’s offspring who carry copies of the gene, is derivative – it is only in virtue of the organism having a given fitness, that it makes sense to talk about the gene’s fitness. So although the formal models may be isomorphic, the biological explanation of why the fitnesses values are as they are, must very different in the two cases. This is why the existence of ‘group effects’ on individual fitness is theoretically interesting, while the existence of ‘genotypic effects’ on genic fitness is not. Hence the technique best suited to detecting such group effects – contextual analysis – is of little interest when applied to diploid population genetics models. This is why the contextual approach produces implausible results when frameshifted downwards.

If the foregoing is correct, it suggests that the role of frameshifting arguments in the levels of selection debate may need re-thinking. To deny any role for frameshifting would presumably be wrong – since the principle of natural selection can be formulated in wholly abstract terms, there is every reason to expect selection processes at different

hierarchical levels to be importantly similar to each other, in some respects. Furthermore, as recent work on the ‘major transitions’ has shown, evolutionary transitions from lower to higher-level entities, e.g. from genes to chromosomes, single-celled to multi-cellular organisms, organisms to colonies etc., are thematically quite similar to each other (Maynard Smith and Szathmary [1995], Michod [1999]). So frameshifting cannot be totally wrong. However, the contextual approach to group selection is theoretically well-motivated, but when frameshifted down to the ‘genic versus genotypic’ level, produces absurd results. A detailed analysis of the limits and validity of frameshifting arguments is clearly required, a task I gladly leave for another occasion.

5.2 Additivity

Finally, I turn to the question of additivity. In the 1980s, Wimsatt ([1980]) and Lloyd ([1988]) argued that the key to the levels of selection question lay in whether the variance in fitness at a given level was additive or not; this view was encapsulated in Lloyd’s ‘additivity criterion’. The basic Wimsatt/Lloyd argument was that if all the variance in fitness at a given hierarchical level is additive, then there is no selection at that level – all the selection must be at a lower level. The additivity criterion has been heavily criticized in the literature, both in relation to genic selection (Godfrey-Smith [1992]; Sarkar [1994]), and group selection (Sober and Wilson [1994]); our concern here will be with group selection. I show below that the relevance of additivity is somewhat different, depending on whether we favour the covariance or contextual approaches.

Sober and Wilson ([1994]) argued that additivity is irrelevant to the group selection issue, on the basis of a simple model for the evolution of altruism in which organisms of two types, altruists (A) and selfish (S), are distributed into groups in varying proportions. Within each group, selfish organisms are fitter, but groups containing a higher proportion of altruists are fitter than groups containing a lower proportion. So long as there is variance in group fitness, both the covariance approach and the contextual approach agree that there is a component of group selection, in this model. For groups vary in fitness depending on their proportion of altruists, hence group fitness and group character (‘proportion of altruists’) covary, so b_{wz} is non-zero; and there are ‘group effects’ on individual fitness, hence the partial regression coefficient β_2 is non-zero. But

as Sober and Wilson point out, it is quite possible for group fitness to be a linear function of proportion of A types (and hence S types.) (Indeed linear fitness functions are commonly used in models of this sort; see e.g. Wilson ([1990]), Kerr and Godfrey-Smith ([2002])). But if so, then all the variance in group fitness will be additive – fitness differences between groups will be fully explicable by the differing proportions of A and S types they contain.¹⁶ So the covariance and contextual approaches agree that, *pace* Wimsatt/Lloyd, perfectly additive variance in group fitness is quite compatible with the operation of group selection.

What this means is that the proposition:

if there is group selection, there must be non-additive variance in group fitness (I)

is false according to both the covariance and contextual approaches. But consider the converse proposition:

if there is no group selection, all the variance in group fitness must be additive (II)

Propositions (I) and (II) are both consequences of the Wimsatt/Lloyd additivity criterion. On the covariance approach, proposition (II) is false, but on the contextual approach, it is true. To see this, we continue to assume that groups contain differing proportions of organisms of two types, A and S. But we do not interpret this as a model for the evolution of altruism – A and S are simply abstract types, and no assumptions about their fitnesses are made. As above, we interpret additivity to mean that all the variance in group fitness can be explained by the differing proportions of A and S types that groups contain; less than perfect additivity means that this is not so.

Now consider proposition II. On the covariance approach, there are only two ways in which there can be no group selection: all the groups must have equal fitness, or group fitness and group character must be uncorrelated. If the former, then there *is* no variance in group fitness, additive or otherwise. If the latter, it does not follow that all the variance in group fitness is additive. On the contrary, if group fitness and group character (proportion of A types) are uncorrelated, this means that fitness differences between

¹⁶ In just the same way, in a diploid population genetics model with two alleles A and B at a locus, if the plot of genotype fitness against ‘proportion of A alleles’ is linear, i.e. if $[w_{AA} - w_{AB}] = [w_{AB} - w_{BB}]$, then all the variance in genotypic fitness must be additive, whatever the allelic frequencies. See Godfrey-Smith ([1992]) or Roughgarden ([1979]).

groups *cannot* be fully explained by the differing proportions of A and B types they contain, so all the variance in group fitness cannot be additive. So proposition II is false.

On the contextual approach, if there is no group selection, this means that there are no group effects on individual fitness – an individual's fitness is a function of its own phenotype alone. This means that an A organism has identical fitness whichever group it is found in, and similarly for a B organism. This immediately implies that all the variance in group fitness *is* additive. For it follows that group fitness must be a linear function of proportion of A types (and hence S types). That is, if we let $w(x)$ denote the fitness of a group containing x A types and $(n-x)$ S types, where n is group size, then $w(x)$ is linear in x – for $[w(x+1) - w(x)] = [W_A - W_S]$, where W_A and W_S are the fitnesses of the A and S types respectively, which are constant by hypothesis. And where group fitness is a linear function of the proportion of A types, it follows that all the variance in group fitness is additive. So if the contextual approach is correct, and the absence of group selection means the absence of group-effects on individual fitness, it follows that if there is no group selection, all the variance in group fitness must be additive – just as proposition (II) says.

To summarize: on the covariance approach, additivity is wholly irrelevant to determining the level at which selection is acting, as Sober and Wilson ([1994]) correctly emphasize. Group selection does not imply that some of the variance in group fitness is non-additive, so (I) is false, and the absence of group selection does not imply that all the variance in group fitness is additive, so (II) is false. On the contextual approach, (I) is also false, but (II) is true - if there is no group selection, then any variance in group fitness must indeed be additive, just as the additivity criterion holds. Of course, this limited result does not rescue the additivity criterion in its full generality, but it does mean that additivity of variance is not *totally* irrelevant to the levels of selection question, just as Wimsatt and Lloyd originally thought.

6 Conclusion

The ongoing controversy over group selection is partially attributable to differences of opinion over what constitutes group selection and what does not. The alternative statistical approaches to group selection explored above reflect such differences. The

simple covariance approach, though plausible at first sight, does not partition the selection differential into components that plausibly correspond to individual and group selection, despite what Price and Hamilton originally thought. The contextual approach does better in this regard. The choice between the contextual approach and the modified covariance approach is harder; I have argued that the former is theoretically preferable, on balance. If this is correct, it means that the role of ‘frameshifting’ arguments in the levels of selection debate needs re-thinking, and that the much criticized ‘additivity criterion’ of Wimsatt and Lloyd may not be entirely wrong, in at least some contexts. What looks like a narrowly technical issue in quantitative genetics turns out to have quite widespread philosophical ramifications.

References

Arnold, A.J. and Fristrup, K. [1982]: 'The theory of evolution by natural selection: a hierarchical expansion', *Paleobiology*, **8**, pp. 113-29.

Boyd, L. H. and Iversen, G. R. [1979]: *Contextual Analysis: Concepts and Statistical Techniques*, Belmont, California: Wadsworth.

Damuth, J. and Heisler, I. L. [1988]: 'Alternative formulations of multi-level selection', *Biology and Philosophy*, **3**, pp. 407-30.

Dawkins, R. [1976]: *The Selfish Gene*, New York: Oxford University Press.

Falk, R. and Sarkar, S. [1992]: 'Harmony from discord', *Biology and Philosophy*, **7**, pp. 463-72.

Frank, S. [1999]: *Foundations of Social Evolution*, Princeton: Princeton University Press.

Godfrey-Smith, P. [1992]: 'Additivity and the units of selection', in D. Hull, M. Forbes and K. Okruhlik (eds.) *PSA 1992*, Vol 1, East Lansing: Philosophy of Science Association, pp. 315-28.

Goodnight, C. J., Schwartz, J. M. and Stevens, L. [1997]: 'Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism', *American Naturalist*, **140**, pp. 743-61.

Hamilton, W. D. [1975]: 'Innate social aptitudes in man: an approach from evolutionary genetics', in R. Fox (ed.) *Biosocial Anthropology*, New York: Wiley.

Heisler, I. L. and Damuth, J. [1987]: 'A method for analysing selection in hierarchically structured populations', *American Naturalist*, **130**, pp. 582-602.

Kerr, B. and Godfrey-Smith, P. [2002]: ‘Individualist and multi-level perspectives on selection in structured populations’, *Biology and Philosophy*, **17**, pp. 477-517.

Lande, R. and Arnold, S. [1983]: ‘The measurement of selection on correlated characters’, *Evolution*, **37**, pp. 1210-26.

Lewontin, R. [1970]: ‘The units of selection’, *Annual Review of Ecology and Systematics*, **1**, 1-18.

Lloyd, E. A. [1988]: *The Structure and Confirmation of Evolutionary Theory*, New York: Greenwood Press.

Maynard Smith, J. and Szathmary. E. [1995]: *The Major Transitions in Evolution*, Oxford: Oxford University Press.

Nunney, L. [1985]: ‘Group Selection, altruism, and structured-deme models’, *American Naturalist*, **126**, pp. 212-30.

Okasha, S. [2001]: ‘Why won’t the group selection controversy go away?’, *British Journal for the Philosophy of Science*, **52**, pp. 25-50.

Okasha, S. [2003]: ‘The concept of group heritability’, *Biology and Philosophy* (forthcoming).

Okasha, S. [2004]: ‘The averaging fallacy’ and the levels of selection’, *Biology and Philosophy* (forthcoming).

Okasha, S. [forthcoming]: ‘Multi-level selection and the partitioning of covariance: a comparison of three approaches’, *Evolution*.

Pearson, K. [1903]: 'Mathematical contributions to the theory of evolution XI. On the influence of natural selection on the variability and correlation of organs', *Philosophical Transactions of the Royal Society of London*, **A200**, pp. 1-66.

Price, G. R. [1970]: 'Selection and covariance', *Nature*, **227**, pp. 529-31.

Price, G. R. [1972]: 'Extension of covariance selection mathematics', *Annals of Human Genetics*, **35**, pp. 485-90.

Robertson, A. [1966]: 'A mathematical model of the culling process in dairy cattle', *Animal Production*, **8**, pp. 93-108.

Roughgarden, J. [1979]: *Theory of Population Genetics and Evolutionary Ecology*, New York: Macmillan.

Sarkar, S. [1994]: 'The additivity of variance and the selection of alleles', in D. Hull, M. Forbes and R. Burian (eds.) *PSA 1994*, Vol 1, East Lansing: Philosophy of Science Association, pp. 3-12.

Sober, E. [1984]: *The Nature of Selection*, Cambridge MA: MIT Press.

Sober, E. and Wilson, D. S. [1994]: 'A critical review of philosophical work on the units of selection problem', *Philosophy of Science*, **61**, 534-55.

Sober, E. and Wilson, D. S. [1998]: *Unto Others. The Evolution and Psychology of Unselfish Behaviour*, Cambridge MA: Harvard University Press.

Tsuji, K. [1995]: 'Reproductive conflicts in the ant *Pristomyrmex Pungens*: contextual analysis and partitioning of covariance', *American Naturalist*, **146**, pp. 586-607.

Wade, M. [1985]: 'Soft selection, hard selection, kin selection, and group selection', *The American Naturalist*, **125**, pp. 61-73.

Wade, M. [1996]: 'Adaptation in sub-divided populations: kin selection and interdemic selection', in M. R. Rose and G. Lauder (eds.) *Adaptation*, San Diego: Academic Press, pp. 381-405.

Wade, M. Goodnight, C. and Stevens, L. [1999]: 'Design and interpretation of experimental studies of interdemic selection: a reply to Getty', *The American Naturalist*, **154**, pp. 599-603.

Walton, D. [1991]: 'The units of selection and the bases of selection', *Philosophy of Science*, **58**, pp. 417-35.

Wilson, D. S. [1975]: 'A theory of group selection', *Proceedings of the National Academy of Sciences USA*, **72**, pp. 143-46.

Wilson, D. S. [1990]: 'Weak altruism, strong group selection', *Oikos*, **59**, pp. 135-40.

Wilson, R. A. [forthcoming]: 'Pluralism, entwinement and the levels of selection', *Philosophy of Science*

Wimsatt, W. [1980]: 'Reductionist research strategies and their biases in the units of selection controversy', in T. Nickles (ed.) *Scientific Discovery: Case Studies*, Dordrecht: Reidel, pp. 213-59.