



Are kin and group selection rivals or friends?

LSE Research Online URL for this paper: <http://eprints.lse.ac.uk/101019/>

Version: Accepted Version

Article:

Birch, Jonathan (2019) Are kin and group selection rivals or friends? *Current Biology*, 29 (11). R433-R438. ISSN 0960-9822

<https://doi.org/10.1016/j.cub.2019.01.065>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Are kin and group selection rivals or friends?

Jonathan Birch

Department of Philosophy, Logic and Scientific Method,
London School of Economics and Political Science,
Houghton Street, London, WC2AE 2AE, UK.

j.birch2@lse.ac.uk
personal.lse.ac.uk/birchj1

For *Current Biology*.

30 October 2018

Abstract: Kin selection and group selection were once seen as competing explanatory hypotheses but now tend to be seen as equivalent ways of describing the same basic idea. Yet this “equivalence thesis” seems not to have brought proponents of kin selection and group selection any closer together. This may be because the equivalence thesis merely shows the equivalence of two statistical formalisms without saying anything about causality. W. D. Hamilton was the first to derive an equivalence result of this type. Yet Hamilton was aware of its limitations, and saw that, while illuminating, it papered over some biologically important distinctions. Attending to these distinctions helps us see where the biological disagreements between proponents of kin selection and group selection really lie.

Once upon a time, kin selection and group selection seemed like competing explanations for the evolution of social behaviour. John Maynard Smith, in the 1964 article in which he coined the term “kin selection”, set out the difference like this:

By kin selection I mean the evolution of characteristics which favour the survival of close relatives of the affected individual, by processes which do not require any discontinuities in population breeding structure. [...] The distinction between kin selection and group selection as here defined is that for kin selection the division of the population into partially isolated breeding populations is a favourable but not essential condition, whereas it is an essential condition for group selection, which depends on the spread of a characteristic to all members of a group by genetic drift. ([1], p. 1145)

Maynard Smith’s exemplar of kin selection was W. D. Hamilton’s original 1964 model, in which organisms are able to act differently towards different categories of close genetic relative (sibling, parent, first cousin, etc.) but in which no discrete groups are posited [2]. His

exemplar of group selection was his own “Haystacks” model, in which discrete, stable groups come to differ genetically from each other due to random drift. These models are strikingly different: they seem to be models of qualitatively different biological processes.

This way of drawing the distinction between kin and group selection has an important advantage: it tells empirical biologists what questions to ask about the cases of social behaviour they care about. *Do the benefits of the behaviour fall on close genetic relatives of the actor? Are partially isolated breeding populations essential to the evolution of the behaviour, or are they inessential?* If the answers are “Yes” and “No”, kin selection is at work. If the answers are “No” and “Yes”, group selection is at work. If the answers are “Yes” and “Yes”, we have an interesting hybrid case in which both processes are at work. This is a great example of how theory can guide and inform empirical biology.

Half a century later, things are no longer so simple, and the connections between theoretical and empirical work on kin and group selection often seem less direct. When we compare today’s social evolution literature to Maynard Smith’s article, the difference that jumps out is that, for many social evolution theorists today, kin selection and group selection (or multi-level selection, as it is now commonly called) are *not competing explanations at all*. They are seen as *two equivalent ways of expressing the same basic idea*. I’ll call this the *equivalence thesis*. It is a point on which key figures in the kin selection tradition, such as Andy Gardner and Stuart West, and key figures in the group selection tradition, such as David Sloan Wilson, seem to agree.

There is an interesting story to tell here about how two ideas which initially seemed so different came to be seen as equivalent (I’ll turn to that below). What’s particularly odd about the current situation is that this broad consensus regarding the equivalence thesis has done very little to neutralize the animosity between the two camps. Proponents of multi-level selection theory use the equivalence thesis to argue that kin selection, although it sounds different, is really just multi-level selection by another name, and they argue that multi-level selection theory captures the biologically important features of the process in a more perspicuous way [3]. Proponents of kin selection theory use the equivalence thesis to argue that any problem analysable using multi-level selection theory can also be analysed using kin selection theory—and they usually proceed to argue that kin selection theory is mathematically richer and better able to handle complexities like class structure [4-7].

As I see it, this failure of the equivalence thesis to reconcile the two camps should lead us to reconsider the equivalence thesis itself. I have come to the view that the equivalence thesis is not a very deep result, biologically speaking. It results from stripping kin selection theory and group selection theory of most of their biological content, and then noting an equivalence between two very abstract statistical formalisms. Somewhere along the line, the literature has lost sight of the importance of articulating *explicitly causal* hypotheses about the evolution of social behaviour. “Kin selection” and “group selection”, if they mean anything, should label causal hypotheses about the drivers of change in a population. And I think both Maynard

Smith and Hamilton were originally correct to think that it's useful to have these separate labels to describe two broad types of causal hypothesis: one that makes a claim about the causal importance of *genealogical kinship*, and one that makes a claim about the causal importance of *discrete and stable groups*. The distinction is not *sharp* by any means, but it's still important. Returning to this older way of drawing the distinction helps us see more clearly the biological issues that are really at stake when theorists from the two camps debate the origins of eusociality and human cooperation.

The quest for generality

How did the equivalence thesis arise? As we've seen, Maynard Smith regarded *reproductively isolated* groups as the mark of group selection. George Price, in what has come to be seen as one of the most important papers on group selection, still assumed the groups to be reproductively isolated [8]. But later work saw a substantial broadening of the concept and a shift in focus towards groups induced by the *interaction structure* of a population rather than its mating structure.

David Sloan Wilson's *trait-group* concept was an important innovation here [9]. A *trait-group* is a group of organisms who affect each other's fitness by expressing a focal trait: this could be a group of beavers making a dam, or a group of insects constructing a nest, or even just a single pair of organisms interacting on a single occasion. In contrast to Maynard Smith's haystacks, trait-groups can be extremely ephemeral, coming into existence and vanishing again many times within a single generation. An organism can participate in many different trait-groups over the course of its life.

At face value, the trait-group concept still implies *discreteness*: it's natural to assume that trait-groups, like haystacks, must have clear boundaries, so that the population can be subdivided into fewer trait-groups than there are individuals. But Wilson, while focussing mainly on discrete groups, also introduced the concept of a *continuous trait-group*, which does not involve positing any discrete boundaries. Each organism's local interaction neighbourhood is its own personal trait-group, and trait-groups blur seamlessly into each other. In this scenario, there are as many trait-groups in the population as there are individual organisms. The advantage of the continuous trait-group concept was that cases of social evolution in viscous populations, with no discernible groups, could be re-described as cases of group selection. The disadvantage was that group selection theory no longer seemed to attach any great significance to the existence, or non-existence, of biologically meaningful groups.

During the same period, kin selection theory was also undergoing a process of generalization. Hamilton had always thought of the "coefficient of relatedness" in inclusive fitness theory as something that was not simply dependent on family trees. He argued that, when selection is weak, it could be *approximated* by Wright's coefficient of relationship, which can be read off family trees. But he thought it should be *defined* in a more general way, as a regression coefficient capturing the genetic similarity between social partners at the genomic loci of

interest. He argued that “kinship should be considered just one way of getting positive regression of genotype in the recipient”, leading him to make the striking claim that “the inclusive-fitness concept is more general than ‘kin selection’” ([10], p. 337). This more general way of thinking about relatedness was promoted by Alan Grafen and David Queller and became orthodox in social evolution research.

Meanwhile, Queller’s “general model for kin selection” also generalized the meaning of “cost” and “benefit” in inclusive fitness theory [11]. Hamilton had attached an explicitly causal meaning to these properties, and saw them as properties of individual social interactions, but they came to be seen as population statistics similar to the coefficient of relatedness: partial regression coefficients in a statistical model of fitness. The advantage was that Hamilton’s famous $rb > c$ rule (see **Box 1**) could now be said to apply to all processes of evolution by natural selection, not just some. The long-running industry of trying to find exceptions to Hamilton’s rule mostly ground to a halt. This allowed Hamilton’s rule to shift from the status of a “rule of thumb” to the status of an organizing framework for the whole field of social evolution research [12]. The disadvantage was that kin selection theory no longer seemed to attach any great significance to the existence, or non-existence, of interactions between genealogical kin.

The disappearance of causality from general theory

These processes of generalization made heavy use of the Price equation, a popular way of representing evolution by natural selection. Price showed that the effects of natural selection can be captured by the *covariance* between the allele of interest and individual fitness. By substituting a regression model of fitness into this covariance, one can very quickly derive Hamilton’s rule, as Hamilton himself showed in 1970 [13]. By breaking up this covariance in a different way, one can very quickly derive a partition of the effects of selection into a “within-group” and a “between-group” component, as Price showed in 1972 ([8], **Box 1**).

Although Price’s work was extremely insightful, one downside to the adoption of his ideas was that, to put it bluntly, causality went missing from general theory in social evolution research [14, 15]. Researchers did not stop thinking about causality, and they didn’t stop making models of particular scenarios based on causal assumptions, but when it came to the *most general statements* of their theoretical commitments, they turned to a purely statistical formalism. Neither the “kin selection partition” of the Price equation nor the “multi-level partition” is explicitly a causal partition, even though both tend to be given a causal *interpretation* [16, 17]. These causal interpretations have always been rather questionable. It’s a general message from the causal inference literature that regression coefficients only quantify causality under very special conditions, and this is still true in the case of social evolution, as recent work by Okasha and Martens has shown [18]. Part of the problem here is that R. A. Fisher called partial regression coefficients “average effects”, and this usage has stuck, even though they only capture *causal* effects under special conditions. They are really statistical notions.

Price's multi-level partition of the Price equation does not even purport to represent causality: it just splits one covariance into two. This led to the development within multi-level selection theory of regression-based methods, under the heading of "contextual analysis" [15], that do purport to represent causality—but these regression-based methods are no more causal than their near-identical counterparts in the kin selection literature.

The statistical nature of the equivalence thesis

The quest for generality, combined with the disappearance of causality from general theory, is what has led to the new consensus that kin selection and group selection are equivalent descriptions of the same process (**Box 1**).

It is a correct result as far as it goes: it is true that the kin selection partition and the multi-level partition of the Price equation can't disagree about the direction and magnitude of change. They split the change into different statistical components, but either approach can be applied correctly in any population. Price's multi-level partition *looks* as though it assumes discrete groups, so cannot be applied in a population without discrete groups, but no such assumption is in fact needed from a mathematical point of view. The multi-level partition, being purely statistical, is still possible even if we assign organisms to groups arbitrarily. There's no formal requirement that the groups have any biological reality.

However, we are now in a position to see why this is a fairly shallow result from a *biological* point of view. It is only possible because explicitly causal assumptions have been avoided. It does not show kin selection and group selection to be equivalent *causal hypotheses*. It shows that they are equivalent when formulated not as causal hypotheses at all, but rather as *statistical descriptions* of change.

Hamilton's view

Hamilton was the first to note that there is a sense in which kin and group selection are equivalent when formulated using the Price equation. Interestingly, though, he did not take this result to obliterate any useful distinction between kin and group selection. He instead offers a subtler view:

If we insist that group selection is different from kin selection the term should be restricted to situations of assortment definitely not involving kin. But it seems on the whole preferable to retain a more flexible use of terms; to use group selection when groups are clearly in evidence and to qualify with mention of 'kin' (as in the 'kin group' selection referred to by Brown), 'relatedness' or 'low migration' (which is often the cause of relatedness in groups), or else 'assortment', as appropriate. The term 'kin selection' appeals most where pedigrees tend to be unbounded and interwoven, as is so often the case with humans. ([10], p. 337)

In the first sentence, Hamilton sounds sceptical of there being a useful distinction to be drawn between kin and group selection. However, what he means is that there is no *sharp*

distinction, or dichotomy. He thinks there is an *unsharp* distinction, admitting of overlap, and he proceeds to set out a nuanced way of thinking about that distinction.

As I read it, he takes the view that there are real and biologically important differences here: it matters whether the population is subdivided into discrete groups, and it matters whether genetic correlations between social partners are explained by genealogical kin interacting preferentially or by something else, such as a greenbeard effect. But the differences that matter are *differences of degree* in aspects of population structure. The degree to which groups are “clearly in evidence” matters, as does the degree to which genetic correlation is explained by kinship as opposed to other causes, but the distinction is not clean or neat. This seems to me to be exactly the right way to think about kin and group selection, and it should guide us in how we use these labels today.

Thinking about kin and group selection using “*K-G* space”

In recent work, I have tried to update and improve on Hamilton’s proposal by finding ways of quantifying the degree to which groups are “clearly in evidence” in a population and the degree to which the genetic assortment is explained by genealogical kin interacting preferentially [12, 19].

Concepts from network theory, such as the “clustering coefficient” and the “relative density” [20], can help us quantify, at any particular moment, the “groupiness” of a social network—the extent to which it contains real, non-arbitrary social groups at that time. If we choose an appropriate measure and take a time-average for the whole population over one generation, we have a rough measure of the extent to which groups are “clearly in evidence”. We can define a quantity G that takes the value 1 when social groups are fully discrete and isolated from each other for long periods (as in the Haystacks model) and the value 0 when there is no population structure at all, with more realistic cases in between.

Meanwhile, the extent to which genetic correlation is explained by kinship can be quantified by comparing the locus-specific correlation with respect to gene of interest to the average correlation across the entire genome, since only kinship can generate correlation at every locus. We can define a quantity K that takes the value 1 when all the correlation is whole-genome correlation and 0 when all the correlation is specific to the locus in question.

Empirical studies of social evolution tend not to measure K or G , but they seem to me to be variables that are worth trying to measure. Both are of intrinsic interest in their own right, but they also help us draw inferences about the sort of social phenomena that are likely or unlikely to evolve in the population.

K matters because it is a guide to how *stable* any altruism that evolves is likely to be. Kinship-based sources of assortment are more likely to lead to stable altruism because they lead to correlation at every locus in the genome; kinship-independent sources, such as

greenbeards and horizontal gene transfer, lead to locus-specific correlations and less stable altruism [21].

By contrast, G matters because it provides some insight into whether or not the population might be undergoing, or about to undergo, an *evolutionary transition in individuality* [22]. All transitions require well-defined groups: in a transition, stable, well-bounded groups of entities at level n become individuals at level $n+1$. “High G ” populations that are also “high K ” are candidates for a *fraternal transition*, in which a reproductive division of labour arises through kin-based altruism [23]. The evolution of multicellularity is a credible example of this. “High G ” populations that are “low K ” may be candidates for an *egalitarian transition*, driven by increasingly close mutualistic cooperation between unrelated entities [23].

These variables lead naturally to the representational device of “ K - G space”. A population’s place in K - G space depends on the extent to which real groups are clearly in evidence and the extent to which genetic correlation is explained by kinship. As Hamilton himself said in other words, selection in high K , low G populations seems aptly described as “kin selection”, whereas selection in high G , low K populations seems aptly described as “group selection”. In the high K , high G region we have hybrid cases that are aptly described as “kin-group selection”, because assortment is kin-based *and* groups are clearly in evidence. In these cases, there really is no meaningful debate to be had about which process is at work.

In the low K , low G region we have various interesting processes of social evolution (such as selection driven by greenbeard effects and by horizontal gene transfer) that it would be misleading to call either “kin selection” or “group selection”. Figure 1 illustrates the general idea and labels four broad regions of the space. It is possible in principle to locate a population exactly in K - G space, given a choice of measure of K and G and information about genetic correlations and interaction structure.

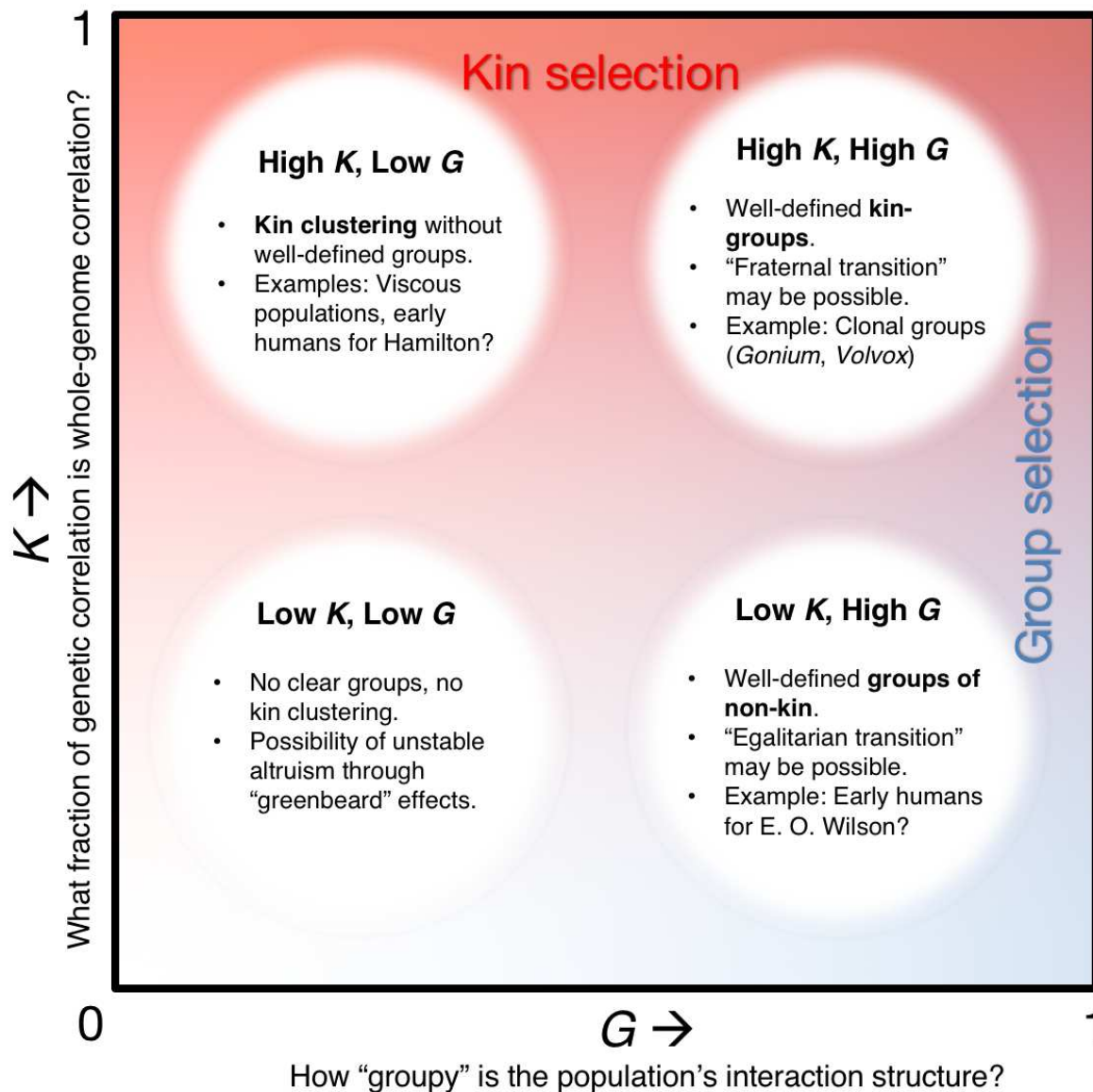


Figure 1: K - G space. “Kin selection” and “group selection” can be visualized as overlapping regions of the space defined by K and G [12].

Why is K - G space useful?

What’s the use of K - G space? It is an unorthodox way of thinking about the relation between kin and group selection, so there had better be some payoff for adopting this unorthodox way of thinking. Otherwise it will just lead to confusion.

The payoff, in my view, is that this representational tool helps us see what is really at stake when proponents of kin selection and group selection debate particular cases, such as the origins of eusociality or the evolution of human cooperation. These are not just non-empirical debates about which mathematical formalism we should use to describe the process. But nor

are they black-and-white clashes between vastly different alternatives. They tend to be debates about where the population of interest should be located in K - G space.

Consider eusociality. Obviously, once eusociality has evolved, groups are clearly in evidence. But there is an empirical debate to be had about whether high G populations are needed for the *origin* of eusociality, or whether the starting point is a fluid and flexible population structure in which insects aggregate facultatively, which only later transitions into a population of well-bounded nests that the workers cannot leave.

Moreover, once eusociality has evolved, kinship-based assortment clearly exists. But there is an empirical debate to be had about whether the origin of eusociality was based on kinship or on a kinship-independent, greenbeard-like source of assortment. When E. O. Wilson and Bert Hölldobler write “what counts is the common possession of eusociality alleles, not relatedness” ([24], p. 13368), they seem to be suggesting, controversially, that assortment was initially *not* kinship-based (and using “relatedness” in a narrower sense than Hamilton’s). Their hypothesis, in short, is that eusociality originated in low K populations (their views about G are somewhat harder to interpret).

Now consider humans. Hamilton made the intriguing observation that human pedigrees tend to be “unbounded and interwoven”, which I take to mean “not such as to form stable, well-bounded kin groups”. This foreshadows recent work on the structure of hunter-gatherer populations, which has pointed to the great fluidity of population structure, with frequent migration between residential camps [25]. If early human populations were similarly fluid, then early human populations were not high K , high G populations.

What, then, drove the evolution of altruism in these populations? One possibility is: intense competition between tribal groups mostly consisting of non-kin, combined with repression of competition within groups. This is roughly the hypothesis of Samuel Bowles and Herbert Gintis in *A Cooperative Species* [26], and it amounts to the claim that early human populations were low K , high G at the tribal level. E. O. Wilson also defends a view like this in *The Social Conquest of Earth* [27]. What all such hypotheses have in common is that they must have story to tell about how competition within groups was repressed, because otherwise the default assumption is that selection within groups will overpower selection between them.

A contrasting causal hypothesis is the following: despite the fluid population structure, people were still able to seek out and preferentially interact with kin. Each individual, while cooperating with campmates for mutual benefit, had their own kinship network spanning multiple camps, and genuinely altruistic behaviour mostly occurred within kinship networks, not across them. This is, in effect, the hypothesis that early human populations were high K , low G . Studies of contemporary hunter-gatherers show that interactions between kin from different camps are more frequent than interactions between non-kin [28]. This is hardly surprising, but it suggests the high K , low G hypothesis merits further exploration.

These examples show how *K-G* space can help us articulate and compare empirical hypotheses. Like Maynard Smith's original proposal, it points empirical biologists towards the questions that matter, though the questions are not quite the same as Maynard Smith's: *To what extent is genetic correlation between social partners explained by kinship? Are groups clearly in evidence, and how stable and well-bounded are they?*

We need to move away from the 1960s view of kin selection and group selection as wholly different processes. But theorists also need to move away from insisting that, because these concepts are equivalent when viewed as statistical partitions of change, one of them can be dispensed with altogether. This has led to a stalemate in which theorists from the two camps continue to disagree, but without being clear as to where the disagreement lies. We should hold on to both terms as useful labels for overlapping regions of *K-G* space, while being pluralistic about the methods we use to analyse the processes in those regions.

Acknowledgements

I thank Samir Okasha, Andy Gardner and Geoff North for their comments on a draft.

References

1. Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, *200*, 1145–1147.
2. Hamilton, W. D. (1964). The genetical evolution of social behaviour I and II. *J. Theor. Biol.*, *7*, 1–52.
3. Wilson, D. S. and Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.*, *82*, 327–348.
4. Lehmann, L., Keller, L., West, S., and Roze, D. (2007). Group selection and kin selection: two concepts but one process. *Proc. Natl. Acad. Sci. USA*, *104*, 6736–6739.
5. West, S. A., Griffin, A. S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.*, *20*, 415–432.
6. West, S. A., Griffin, A. S., and Gardner, A. (2008). Social semantics: how useful has group selection been? *J. Evol. Biol.*, *21*, 374–383.
7. Gardner, A. (2015). The genetical theory of multi-level selection. *J. Evol. Biol.*, *28*, 305–319.
8. Price, G. R. (1972). Extension of covariance selection mathematics. *Ann. Hum. Genet.*, *35*, 485–490.
9. Wilson, D. S. (1975). 'A Theory of Group Selection', *Proc. Natl. Acad. Sci. USA*, *3*, 143–146.
10. Hamilton, W. D. (1975). Innate social aptitudes of man: an approach from evolutionary genetics. In Fox, R., editor, *Biosocial Anthropology* (New York: Wiley), pp. 133–55.
11. Queller, D. C. (1992). A general model for kin selection. *Evolution*, *46*, 376–80.

12. Birch, J. (2017). *The Philosophy of Social Evolution* (Oxford: Oxford University Press).
13. Hamilton, W. D. (1970). Selfish and spiteful behaviour in an evolutionary model. *Nature*, 228, 1218–1220.
14. Birch, J. and Okasha, S. (2015). Kin selection and its critics. *BioScience*, 65, 22–32.
15. Okasha, S. (2016). The relation between kin and multi-level selection: An approach using causal graphs. *Br. J. Philos. Sci.*, 67, 435–470.
16. Okasha, S. (2006). *Evolution and the Levels of Selection* (Oxford: Oxford University Press).
17. Birch, J. (2014) Hamilton’s rule and its discontents. *Br. J. Philos. Sci.*, 65, 381–411.
18. Okasha, S. and Martens, J. (2016). The causal meaning of Hamilton’s rule. *R. Soc. Open Sci.*, 3, 160037.
19. Birch, J. (2018). Kin selection, group selection, and the varieties of population structure. *Br. J. Philos. Sci.* Published online. <https://doi.org/10.1093/bjps/axx028>
20. Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486, 75–174.
21. Gardner, A. and West, S. A. (2010). Greenbeards. *Evolution*, 64, 25–38.
22. Michod, R. E. (1999). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality* (Princeton, NJ: Princeton University Press).
23. Queller, D. C. (2000). Relatedness and the fraternal major transitions. *Phil. Trans. R. Soc. Lond. B: Biol. Sci.*, 355, 1647–1655.
24. Wilson, E. O. and Hölldobler, B. (2005). Eusociality: Origin and consequences. *Proc. Natl. Acad. Sci. USA*, 102, 13367–13371.
25. Hill, K. R., Walker, R. S., Božičević, M., Eder, J., Headland, T., Hewlett, B., Hurtado, A. M., Marlowe, F., Wiessner, P., and Wood, B. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science*, 331, 1286–1289.
26. Bowles, S. and Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and its Evolution* (Princeton, NJ: Princeton University Press).
27. Wilson, E. O. (2012). *The Social Conquest of Earth* (New York: W. W. Norton & Co.).
28. Hill, K. R., Wood, B. M., Baggio, J., Hurtado, A. M., and Boyd, R. T. (2014) Hunter-gatherer inter-band interaction rates: implications for cumulative culture. *PLoS ONE*, 9, e102806.
29. Marshall, J. A. R. (2011). Group selection and kin selection: formally equivalent approaches. *Trends Ecol. Evol.* 26, 325–332.

Box 1: The equivalence thesis

The best-known argument for the ‘equivalence’ of kin and group selection involves comparing a generalized version of Hamilton’s rule [11] with the multi-level version of the Price equation [8] and noting that both provide correct conditions for positive gene frequency

change in a group-structured population given a small number of assumptions (for details, see [14, 29]).

The route to the generalized version of Hamilton's rule involves mathematically decomposing the overall change in allele frequency between generations into a 'direct' and an 'indirect' component:

$$\Delta\bar{p} = \overbrace{\frac{-c\text{Var}(p_i)}{\bar{w}}}_{\text{direct fitness effects}} + \overbrace{\frac{rb\text{Var}(p_i)}{\bar{w}}}_{\text{indirect fitness effects}} \quad (\text{Kin selection partition})$$

Here, p_i denotes the individual gene frequency, with respect to the allele of interest, of the i th individual in the population (i.e. its number of copies of the allele divided by its ploidy); and $\text{Var}(p_i)$ is the variance of p_i in the ancestral population. $\Delta\bar{p}$ is the change in the frequency of the allele between the ancestral and descendant populations, and \bar{w} is the mean fitness in the ancestral population. r is the coefficient of relatedness, defined as a measure of the statistical association between the genotypes of social partners. c and b are, respectively, the coefficients of cost and benefit, defined as partial regression coefficients in a regression model of fitness.

The above decomposition implies the following condition for positive change:

$$\Delta\bar{p} > 0 \iff rb > c, \text{ provided } \text{Var}(p_i) \neq 0 \quad (\text{Hamilton's rule})$$

Now compare this with the multi-level Price equation (Price [1972]). This provides an alternative mathematical decomposition of change that applies whenever organisms are sorted into groups. It partitions change into a 'between-group' and 'within-group' component:

$$\Delta\bar{p} = \overbrace{\frac{\text{Cov}(W_k, P_k)}{\bar{w}}}_{\text{between-group}} + \overbrace{\frac{E_k[\text{Cov}_k(w_{jk}, p_{jk})]}{\bar{w}}}_{\text{within-group}} \quad (\text{Multi-level partition})$$

Here, w_{jk} and p_{jk} denote the fitness and individual gene frequency (respectively) of the j th member of the k th group; while W_k and P_k denote (respectively) the mean fitness and group gene frequency (respectively) of the k th group. $\text{Cov}(W_k, P_k)$ captures the covariance between a group's gene frequency and its mean fitness, while $E_k[\text{Cov}_k(w_{jk}, p_{jk})]$ captures the average across groups of the within-group covariance between an individual's gene frequency and its fitness.

This decomposition also implies a condition for positive change, which we might call 'Price's rule':

$$\Delta \bar{p} > 0 \iff \text{Cov}(W_k, P_k) + E_k[\text{Cov}_k(w_{jk}, p_{jk})] > 0 \quad (\text{Price's rule})$$

The argument for the ‘equivalence’ of the two conditions relies on the fact that both are derived from the Price equation (Price [1970]) with few additional assumptions. In all populations that satisfy the assumptions of both derivations (that is, in all group-structured populations in which alleles are transmitted without bias, c and b are well-defined, and $\text{Var}(p_i) \neq 0$) both decompositions are correct and the following equivalence holds:

$$\Delta \bar{p} > 0 \iff rb > c \iff \text{Cov}(W_k, P_k) + E_k[\text{Cov}_k(w_{jk}, p_{jk})] > 0 \quad (\text{Equivalence thesis})$$

To understand the intuitive rationale for the equivalence thesis, imagine the typical circumstances under which each condition is satisfied for an altruistic trait controlled by a single gene. First, consider what is required for $rb > c$. It must be that bearers of the gene cluster together, so that the benefits of altruism fall differentially on other bearers of the gene. Second, consider what is required for the selection against the trait within groups to be outweighed by selection for the trait between groups. Again, it must be that bearers of the gene cluster together, so that the heritable variation in fitness within groups is suppressed and the heritable variation in fitness between groups is boosted. Hamilton’s rule and Price’s rule are alternative ways of capturing the fundamental requirement that there is correlated interaction between bearers of the gene for altruism.
