

# Group Selection: A Review Essay on *Does Altruism Exist?* by David Sloan Wilson

ARTHUR J. ROBSON\*

December 24, 2015

## 1 Introduction

David Sloan Wilson (2015) argues forcefully that altruism exists in humans and that the biological mechanism of “group selection” is responsible. Indeed, Wilson argues that altruism and group selection are pervasive and important across non-human species too, contrary to the orthodox view in biology. This review focusses on the implications of Wilson’s argument for humans and for economics, in particular. Full appreciation of the human situation can only be obtained, however, by also considering other species, to provide reinforcement or contrast.

At stake here is the most fruitful view of the basic nature of human beings. As an economist, one is trained to be skeptical of the need to suppose that individuals are motivated by the common good. Economic theory has done well in explaining a wide range of phenomena on the basis of selfish preferences, and the related orthodox view within biology of the individual as the unit of selection is accordingly highly congenial to economists.

There are aspects of human economic behavior that are tempting to explain by group selection. For example, human beings are often willing to trade with strangers they will likely never see again, as might be modelled as cooperation in the one-shot prisoner’s dilemma. However, to make the case that a radical re-vamping of the discipline is necessary requires a comprehensive cataloging and evaluation of the empirical phenomena that contradict the current approach. What exactly is this evidence? Do the most obvious and significant failures of orthodox economics arise from neglecting altruism? Is there no way of accommodating the anomalies within the current orthodox theory? Would a more modest extension of orthodox economics than that implied by group selection be sufficient?

---

\*Simon Fraser University

Despite the forensic skill that Wilson has, it is a weakness of his argument that he does not present compelling empirical evidence. Much of the evidence he presents can be reinterpreted within the orthodox theory. In general, the importance of any evidence that cannot be accommodated needs to be assessed. It is not enough, in the end, that there exist a few peripheral patterns of behavior that suggest the need for a transformation.

A key feature of economics is that at its heart there is a simple and general model. Maintaining this is scientifically appropriate if there are only occasional relatively minor empirical deviations. Even if the empirical anomalies within economics are ultimately found to be pervasive and important, a complete re-vamping of economics requires the development of a viable alternative. It seems desirable that this alternative retain the simplicity and generality of the current theory.

The general argument against group selection, for non-human species, is not that it is, in any sense, logically flawed; rather it is widely believed within biology that the required parameter values are simply empirically implausible. It is true, as shown by Bowles in work that is discussed below, that this quantitative case against group selection is weakened for humans given the importance of human culture. Nevertheless, it seems premature to embrace a general theory of economics founded on group selection. It is not even yet clear what conventional evolutionary theory based on the individual as the unit of selection implies for economics. I argue below that biological evolution based on individual selection is indeed a useful basis for economics, one that will preserve much but not all of current theory. It is less clear what exactly group selection would imply for economics. However, it does seem likely that introducing group selection would considerably reduce the empirical predictions of economics.

It seems reasonable, then, that the implications for economics of orthodox individual selection should be completely derived before contemplating the likely radical implications of group selection.

**Outline of this Review** This review is structured as follows. First, in Section 2, I present a little useful background on the highly charged debate in biology concerning the appropriate level of selection—should it be that of the group, the individual, or the gene?

Then, in Section 3, I discuss Wilson’s thesis in the context of relevant recent literature on this issue. First, I discuss how the picture of humans painted by Wilson in Chapter 4 is a stark one where altruism would be the norm and selfishness, let alone malevolence, the exception. I contrast the high level of altruism that this would imply with key evidence that Wilson uses to illustrate what altruism means in practice, in Chapter 8. This evidence of more modest altruism is largely consistent with orthodox economics. I then focus on aspects of Wilson’s thesis that seem important for humans in general and economics in particular—how culture improves the plausibility of group selection for humans and how culture and genes “co-evolve”. Next, I critically evaluate influential work done by Wilson (and by Sober and Wilson) that buttresses group selection

by drawing a theoretical parallel to kin selection. Finally in this section, I discuss work by Bowles that increases the plausibility of group selection for humans by using data for hunter-gatherers.

On the one hand, to show the usefulness of orthodox evolutionary theory, I then consider in Section 4 the implications of fitness as a relative rather than absolute biological criterion. In Chapter 2, Wilson eloquently argues for this orthodox view in biology in favor of relative fitness. I sketch how this doctrine sheds light on phenomena in economics that seem puzzling from a conventional perspective—a greater aversion to aggregate rather than idiosyncratic risk, for example.

On the other hand, to illustrate a key drawback of embracing group selection, in Section 5, I finally apply a famous model of group selection—the haystack model. This shows how an unknown intensity of group selection will enlarge the set of possible outcomes, thereby reducing the falsifiability of the generalized theory.

## 2 Background

For the sake of non-biologists, it is worth sketching the history of the debate concerning individual and group selection—a highly charged debate that lies at the heart of evolutionary biology.

**Group selection** Group selection stands largely in contrast to individual selection. Individual selection is illustrated by considering an unstructured asexually reproducing population in which there are a number of variants. These variants have differing suitability to the environment, or fitness, as reflected in a varying number of offspring. In general, the environment is affected by the choices of other individuals. Consider, for example, the prisoner’s dilemma. The choice of cooperate is “altruistic” since it raises the payoff of one’s opponent, but is costly to oneself. For any possible distribution of the population between the defect and cooperate strategies, the choice of defect does better than the choice of cooperate. Not surprisingly, then, as time goes by, the proportion of defectors increases in the population, eventually converging to one.

Suppose now that the population is divided into groups. Within a group that has both defectors and cooperators, the defectors will grow, as a fraction of the group. If there are groups composed only of cooperators, however, these groups will grow faster than the groups composed mostly of defectors. The evolutionary outcome in this structured population then depends on the details of how the groups are maintained and renewed. In order to obtain cooperation, there needs to be limited migration between groups so that mutant defectors do not infect the groups of cooperators. In addition, the consequences of high group payoffs, as obtained by the cooperators, need to be substantial.

Group selection has intuitive appeal.<sup>1</sup> For example, the argument that individuals age and die in order to leave more food for younger and more fertile

---

<sup>1</sup>The following discussion draws on Robson (2008).

non-relatives involves a naive form of group selection. Such an argument must address why a mutant type that did not age, therefore having more descendants, could not successfully invade a population that aged, even if the final consequence were bad for the species.

A well-known advocate of group selection was Wynne-Edwards (1962) who proposed that birds limit the size of their clutches of eggs so that the total population of the species would not strain the carrying capacity of the environment.

Such group selection arguments were undercut by Williams (1966). If a mutant type of bird has a larger clutch, he argued, this more fertile type would surely take over the population, whatever the consequences for the species. Further, there are good selfish reasons for an individual to limit her clutch size. It is plausible that, beyond a certain point, increasing the clutch size would only reduce the expected number of offspring that survive to reproduce. This is because an increased number of eggs would reduce the parental resources devoted to each one. It would be optimal then for an individual to limit her clutch size.

Dawkins (1976) has been vocal in rejecting group selection, arguing further that it is indeed the gene that should be cast as the unit of selection. In principle, at least, genes might then evolve to the detriment of the individual in the same way that individuals might evolve to the detriment of the group. The mechanism by which genes combine in sexual reproduction and are expressed in an individual surely matters. However, it seems plausible that the interests of the gene and the individual do not conflict in many instances.

In any case, our ignorance of the details of the transformation of genes into individual traits, particularly for complex behavioral characteristics, leaves us little choice other than to consider the level of the individual, as a proxy for the gene. Grafen (1991), for example, advocates this approach, in his so-called “phenotypic gambit.” Hence, despite the theoretical primacy of the gene, it is often unavoidable to limit the comparison to the individual level and the group level of selection. This is Wilson’s approach throughout.

This debate is anticipated in Darwin, which is remarkable given that detailed understanding of genetic inheritance was then lacking. In particular, he emphasized that a certain variation would spread if this variation led to greater reproductive success for *individuals* and was genetically transmitted to their descendants.

However, Darwin (1871) also dabbled in group selection, when considering the implications of his theory for humans. Thus, he thought a group composed of individuals who had qualities that were of military value to the group but costly to the individual would thrive at the expense of groups that did not have such individuals. Thus natural selection would generate humans who engaged in behavior beneficial to the survival of a group, even if this behavior had a fitness cost to the individual.

### 3 Wilson’s Thesis in Context

**A Major Evolutionary Transition?** It sets the stage to recognize up-front the strength of Wilson’s claims for humans. Chapter 7 contains some of Wilson’s critical views of conventional economics. He argues, for example, that economists misrepresent Adam Smith’s overall views by focussing too much on the Invisible Hand from Smith (1776), at the expense of the rather more eclectic picture painted in the *Theory of Moral Sentiments* (Smith, 1761). Purely as a commentary on Smith, this may well be fair.

It is in Chapter 4, however, that Wilson stakes out a diametrically opposite empirical view of humans to that of economics. He does not merely suggest that human altruism might occasionally arise, but argues here that humans have undergone a “major evolutionary transition”. That is, we have reached a status where societies composed of individual humans function as smoothly as an individual organism that is made up of a multitude of individual cells. For an organism, there need not be perfect harmonization of the “interests” of cells, as the existence of phenomena like cancer attest. So he would admit analogous maladaptations of individual humans to society. However, Wilson’s theory still involves a high degree of altruism, and so seems likely to fit observation less well than does current economic theory.

It may not be easy to precisely quantify the degree of harmonization of individuals within human society, but it seems doubtful that we have attained such a high level of social efficiency. For that matter, not only does it seem that people may be benevolent on occasion, contradicting the orthodox economic model, but they are also malevolent on occasion, which equally contradicts the standard model. It is not obvious that benevolence is more pervasive than malevolence. If we are going to incorporate the former shouldn’t we also explain the latter?

It is not out of the question that, in the future, group selection might be a component of a biological basis for economics. To be useful, however, such an extended economic theory would need to embrace group selection in a more limited and conditional sense than Wilson suggests. Possibly, for example, group selection would foster altruism to fellow group members, but hostility to outsiders. Adding unconditional altruism seems likely to contradict the facts; adding a variable but unknown intensity of altruism is bound to reduce the predictive power of economics, as is shown in Section 5 below.

**Evidence of Altruism for Wilson** It is illuminating, however, to examine the observations that Wilson presents as being altruism. The extent of this observed altruism is more modest than Chapter 4 would suggest. Indeed, to an economist, this evidence does not build a strong case for pervasive group selection. There is an issue here concerning the meaning of the term “altruism”. That is, Wilson considers benevolent repeated interactions of an individual and his or her neighbors to be altruistic; economists would mostly interpret this merely as enlightened self-interest.

The discussion of this evidence is in Chapter 8, which reports on a fascinating

empirical project that Wilson has been involved with in Binghamton, NY. This project administered a test of “pro-sociality” to public school students in grades 6-12. This test considered, in particular, attitudes towards others, as measured by their agreement or disagreement with statements like “I think it is important to help other people”. As Wilson realizes, positive answers to such questions might reflect little more than a desire to look good to the investigator. Next, Wilson and his coauthors showed that these scores varied systematically across neighborhoods in Binghamton, with a high correlation between an individual and his or her neighborhood of 0.7. Finally, and crucially, Wilson and his coauthors verified that these scores are correlated with observable behavioral differences. For example, one approach taken was to drop stamped addressed envelopes on the sidewalks in various neighborhoods to obtain the rate at which such envelopes were mailed back. The test scores on pro-sociality did help predict actual cooperative behavior, although the precise correlations are not reported.

As Wilson discusses, this is a notion of altruism that adapts readily to the circumstances. What is it exactly that is generated here by group selection? The behavioral flexibility shown means that it is not purely a genetically programmed urge to be benevolent unconditionally. Possibly there might be a genetically driven urge to anticipate and reflect the ambient behavior in one’s neighborhood.

Although it is not possible to dismiss genetic effects completely here, standard game theory, with no altruism at all, would explain these observations quite well. Consider the repeated prisoner’s dilemma with a discount factor close to one. There are a multitude of perfect equilibria—defect in all periods whatever the history being one of them. It is easy to support full cooperation also. If you know are in an environment where defect is always played, you will defect right away and always thereafter. If you know you are in an environment where a fully cooperative equilibrium is played, you will adhere to the appropriate strategy.

There are admittedly loose ends in the game theoretic account—for example, people who mail back stamped addressed envelopes have no realistic expectation of a payback directly caused by that action. A plausible reason that cooperation might depend on the environment only in a rough fashion that leaves such loose ends, is that it involves a rule of thumb. This rule of thumb might be directly genetically programmed or, more plausibly, it might be generated on the spot by flexible but limited intelligence. All of these issues are worthy of further investigation. In the end, however, although the evidence here is fascinating, it does not seem compelling evidence of the need to allow group selection and altruism, in a strong sense that exceeds enlightened self-interest.

**The Improved Plausibility of Group Selection with Culture** The conventional wisdom among biologists is that group selection is theoretically possible. On other hand, it is believed that plausible values of the various parameters make group selection too weak relative to individual selection to matter much. That is, for purely genetic transmission of characteristics, it seems that there is

too much migration between groups for the groups to persist long enough to be subject to selection. Further, it also seems that selection against groups that make choices that are group suboptimal would not be severe enough.

These arguments against group selection are markedly less compelling in the case of human beings. This was admitted by Ernst Mayr (1991), for example, despite his opposition to group selection for other species. He noted that the crucial relevant aspect of human beings is the huge role played by culture. Culture is an alternative pathway by which behavior can be transmitted. Compared to genetic inheritance, culture is very flexible, and transmitted at a much faster rate.

Indeed, it is clear that social groups exist within humans that are defined using cultural mechanisms—language, for example. These groups are potentially rather persistent, with the ability to limit entry or exit. This feature contrasts with purely genetic evolution involving groups in non-human species, where the migration of individuals between groups often implies that there is too little variation across groups for natural selection to be very effective.

At the same time, culturally defined groups that attain lower payoffs than others may be subject to rapid effective extinction. Such extinction need not entail the demise of individuals, but might correspond merely to a less successful group being sensitive to the greater success of another group's choices and so adopting these choices. Alternatively, members of the less successful group might be absorbed into the more successful one and so acquire its behavioral characteristics. The plausibility of group selection for humans is then an empirical question.

**Coevolution of Genes and Culture** In Chapter 5, Wilson considers further how the evolution of genes and culture might be intertwined, how they might “coevolve”, that is. Genes and culture might coevolve in various ways. Most simply, perhaps, our propensity to imitate others, which is the basis for culture in the present sense, could be a mechanism to generate behavioral flexibility. It is this propensity that might be genetically encoded and so generated by natural selection.

Robalino and Robson (2013), consider a variant of this argument. An individual has a deep payoff function that is simply biological fitness. This deep payoff depends on money outcomes, but also on an unknown state. A signal, which is culturally transmitted, changes the belief distribution over the state. Such a shift would apparently change the preferences over money outcomes. Hence, for example, risk-aversion might apparently increase in response to a disaster.

On the other hand, culture might also drive genetic evolution. For example, the introduction of agriculture created an environment in which lactose intolerance in adults was actively selected against. That is, culture created an environment that favored certain genotypes over others. (Boyd and Richerson, 1985, is a key treatise on such two-way interactions—the “co-evolution” of genes and culture.)

It seems desirable for Wilson to clarify, or at least narrow down, how a preference for altruism might be genetic in some ways and cultural in others. What aspects of behavior are cultural and hence subject to rapid modification? What aspects of preferences or behavior are hard-wired and hence less flexible? Specifying this decomposition would increase the empirical content of the theory. Given the plasticity seen in the data, a preference for altruism seems unlikely to be purely genetic.

If altruism arose in small hunter-gatherer groups, how is it triggered if it now arises in a much larger community? How is it that we now treat strangers we encounter in daily life in some respects almost as if they were members of the same small group? Perhaps we simultaneously evolved a willingness to treat outsiders to the group in a hostile fashion. If so, why don't we use the "defect against strangers" response nearly always in modern circumstances?

**Kin Selection as Group Selection** Group selection has recently undergone a partial renaissance in biology, a renaissance to which Wilson has contributed substantially. Wilson has demonstrated the theoretical coherence of group selection. Sober and Wilson (1999), for example, show this by making the case that group selection is implicit in "kin selection". Kin selection is the biological theory that explains why a mother, for example, would make material sacrifices that benefit her offspring. That this is true, in fact, is evident and the biological explanation is highly satisfactory.

This explanation is that a rare gene variant that promoted such sacrifice in the mother would arise with probability  $1/2$  in each offspring, given the detailed mechanics of sexual reproduction. If the mother makes a sacrifice in favor of her offspring, it is as if the rare gene were doing itself a favor with probability  $1/2$ . Not merely does this theory predict maternal benevolence, but it precisely quantifies it.

Wilson recapitulates this argument from Sober and Wilson in Chapter 3. Here, he describes a notion of how two theories might be "equivalent"—a notion of isomorphism under which either might be accepted as an equally attractive account of the facts. He argues that, in particular, kin selection is equivalent to group selection. Wilson's claim that kin selection is theoretically isomorphic to group selection seems valid. However, this observation by itself does not empirically validate group selection in general.

What group selection requires is limited migration between the groups. It is this indeed that gives the groups their meaning. If we think of kin selection as group selection, the groups are formed of the offspring of a particular female who are being raised together by her. This is a compelling mode of group formation that would induce very limited mobility.

To apply group selection without relying on kinship requires that there exist a mechanism that promotes group cohesion to a comparable extent to that induced by kinship. Consider, for example, the maintenance of cooperation in the prisoner's dilemma. Suppose, for simplicity, this would be the outcome under kin selection when the game is played between two full relatives, individuals



who are clones of one another. This kin selection mechanism works because a mutant that chose defect instead of cooperate would face a sibling who is a fellow mutant for sure and therefore this mutant will do less well than the original cooperators.

For this to work under general group selection, there needs to be some device that forces mutant defectors to preferentially face one another rather than cooperators, even when there are only a small number of such defect mutants. What if the mutants send a signal of their mutant status, and use that to find fellow mutants? This would not function as a suitable device, in the end, however, since a new mutant that failed to signal, or ignored the signal, and so exploited cooperators, would outperform the original mutant and the cooperators.

In the final analysis, the theoretical isomorphism of kin selection to group selection does not guarantee that the latter is a persuasive theory in the absence of groups composed of close kin. The plausibility of group selection hinges instead on a factual matter: What mechanisms exist that maintain groups and reward successful groups with increased representation in the population?

**Estimating the Strength of Group Selection for Humans** Bowles addresses this question for humans. In Bowles (2006, 2009), he presents a meticulous and painstaking argument for the empirical possibility of group selection, using evidence from our hunter-gatherer past. His argument relies on intergroup warfare, which was particularly lethal in humans, and on there being a human population that is estimated to have been sufficiently varied genetically. Also required here for group selection is culturally transmitted “reproductive leveling”, which includes distinctively human practices such as food sharing and monogamy. “Reproductive leveling” describes the reduced advantage that is imposed socially on individuals who do relatively well. This reduces the force of individual selection within a group, while the lethal nature of warfare increases the force of group selection. Bowles’s model naturally entails hostility to outsiders in addition to altruism to insiders. This then produces a glimmering of an explanation for the possibility of antagonism.

The theory that Bowles favors could readily generate empirical predictions that differ sharply from those obtained from orthodox evolutionary theory. For example, the latter generates defection in the one-shot prisoner’s dilemma; the former could well generate cooperation. The implications of a variable intensity of group selection are considered below in Section 5.

Bowles’s analysis provokes the questions: What accounts for the reproductive levelling in humans? Why is it evolutionarily stable? Could a mutant that somehow opted out of such sharing outperform the levellers and hence invade a group practicing levelling?

Bowles makes group selection in humans a more plausible possibility. His argument is much more nuanced than Wilson’s, however. In the first place, Bowles’s argument does not generate unconditional human altruism. It encompasses hostility to outsiders, for example. Neither does Bowles’s argument extend to other species, given that these results cannot be obtained without

some degree of reproductive levelling.

Even in the case of humans, establishing the possibility of group selection leaves open a number of difficult and crucial issues. How long have the hunter-gatherer circumstances considered here plausibly been in effect? Wouldn't our evolutionary history previous to that, especially that of ancestral non-human species, still cast a long shadow, implying a core element of selfishness? What exactly is the new theory of economics that adding an element of group selection generates? When and why are we altruistic? If, for example, this theory implies cooperation within groups and hostility to outsiders, how did our perception of the group grow to encompass many of the unfamiliar individuals that we currently interact with? Would there be exceptions? Surely the stock traders on the floor of a particular exchange are not driven much by altruism, for example, despite the familiarity of traders to one another?

## 4 New Implications of Individual Selection for Economics

In this section, I consider some new implications of basing economics on orthodox evolutionary theory in which the individual is the unit of selection. The most fruitful approach might then well be to consider the full implications of such evolutionary theory before resorting to group selection. This is especially so since group selection may well weaken the empirical predictions of economics substantially. This issue is taken up in the next section.

**Relative and Absolute Fitness** In Chapter 2, Wilson addresses the primacy of relative fitness over absolute fitness in determining whether a particular “phenotype” will emerge as evolutionarily dominant. This has important ramifications for evolved economic preferences—in particular, these may well be different than the preferences that are standard in economics.

It is not that absolute utility is always incorrect, however, since, in many situations excelling in a relative sense is equivalent to excelling in an absolute sense. In a standard evolutionary game setting with a large population, for example, where individuals are paired to play a symmetric game, it is enough to maximize one's own payoff without worrying about that of your opponent. Only if the population is small might it pay to lower one's opponents payoff. That is, if the population is large your relative fitness is equivalent to your absolute fitness, since your particular current opponent's payoff does not affect the *average* of all other individuals. Only if the population is small would it pay to lower the payoff of your particular current opponent, since lowering this will then have an appreciable effect on the average payoffs of all other individuals (see Schaffer, 1988, for example).

To derive attitudes to risk, on the other hand, it may be necessary to consider relative fitness. Suppose that individuals face a menu of possible gambles over offspring, where all of the risk involved is idiosyncratic, that is, independent

across individuals. Then the most successful type will be the one that maximizes expected fitness, which is effectively absolute.

Suppose, however, the gambles over offspring are aggregate, so that individuals obtain the same realizations. Then Robson (1996a) shows that relative fitness will play a irreducible role (see also Robson and Samuelson, 2011). Consider the following example concerning a (stylized) Arctic hare in the Canadian North. (This example is inspired by Cooper and Kaplan, 1982.) The probability of a snowy winter is  $p \in (0, 1/2)$ ; whereas the probability of a clear winter is  $1 - p \in (1/2, 1)$ , and these winters are drawn independently each year. Hares that keep dark coats survive clear winters for sure but die in snowy winters; those that develop white coats survive snowy winters but die in clear ones. The nature of the winter constitutes an aggregate shock. Hares of all types each have  $f$  offspring in the spring, if they survive. Either pure type will die out completely on the first occurrence of an unfavorable kind of winter.

Consider then a type that randomizes, where individuals of this type independently choose a white coat with probability  $\pi$  and a dark coat with probability  $1 - \pi$ . After  $T$  years, the number of the randomizing type is then

$$S(T) = \pi^{n(T)}(1-\pi)^{T-n(T)} f^T, \text{ where } n(T) \text{ is the random number of snowy winters drawn in } T \text{ trials.}$$

It follows that

$$\frac{1}{T} \ln S(T) = \frac{n(T)}{T} \ln \pi + \frac{T - n(T)}{T} \ln(1 - \pi) + \ln f,$$

so that, by the strong law of large numbers,

$$\frac{1}{T} \ln S(T) \rightarrow r = p \ln \pi + (1 - p) \ln(1 - \pi) + \ln f, \text{ as } T \rightarrow \infty, \text{ with probability 1.}$$

Indeed,  $r$  is an appropriate measure of the growth rate. Any type with a higher limiting value of  $\frac{1}{T} \ln S(T)$  will swamp a type with a lower limiting value. The most successful type then maximizes  $r$ , as is achieved by choosing  $\pi = p$ .

Cooper and Kaplan (1982) describe this situation in a fashion that might tempt an economist. They say that individuals who choose a white coat after the flip of the coin are “altruistic” because the probability of such an individual dying is higher than the probability of an individual with a dark coat dying since  $1 - p > 1/2 > p$ . But the issue here is the correct notion of fitness, as a biologist would express it, or the correct utility function, as an economist would.

Grafen (1999) argued in response to Cooper and Kaplan that the correct notion of fitness is relative and that the optimal type is indeed maximizing this correct notion of fitness. Consider a continuum of animals of size 1, say. Suppose a fraction  $\pi$  of these choose white and a fraction  $1 - \pi$  choose dark. Now consider the choice of a small mass of individuals of size  $\varepsilon$ . If they choose white, the expected fraction of the population they will constitute the following spring is  $\frac{p\varepsilon}{\pi} = \varepsilon$  if  $p = \pi$ . If they choose dark, the expected fraction of the population they will constitute is  $\frac{1-p}{1-\pi}\varepsilon = \varepsilon$ , if  $p = \pi$ . It follows that the type that randomizes  $(\pi, 1 - \pi)$  is maximizing the expected fraction of the population it will comprise,

so this relative notion is the right notion of fitness. White is equivalent to dark just because the overall survival probability of white individuals is worse than that of dark.

It is not enough to consider just the probability of death, as absolute fitness would imply. It is necessary to consider how well you are doing when you do survive relative to others. Dead is just dead, zero fitness, but when you survive it matters how large you loom in the population, so fitness is relative. Curry (2001) essentially reinterprets Grafen’s point as it applies to utility. The appropriate utility function is not absolute, but relative and hence interdependent.

That fitness is fundamentally relative rather than absolute, as Wilson states, is crucial to understanding evolved economic preferences, and this observation applies with full force when the level of selection is that of the individual.

## 5 The Empirical Implications of Group Selection

What would be the empirical consequences of adding some intensity of group selection to the biological foundation of economic preferences? A crucial methodological issue with adding group selection and hence altruism to our picture of humans is that it may well produce a theory that is less “falsifiable” (Popper, 2002).

To investigate this issue, consider a version of a classic model of group selection, the “haystack” model of Maynard Smith (1964) that has been simplified in several ways.<sup>2</sup> In the first place, all reproduction here is assumed here to be asexual. That is, each female mouse produces offspring that are precise copies of their mother. There are a number of haystacks in a farmer’s field, where each haystack is originally colonized by two mice, as another simplification. Each pair of mice play the prisoner’s dilemma, with the usual two choices—cooperate or defect. Payoffs for each individual take the concrete form of the number of offspring. Offspring inherit their mother’s choice of strategy. There are a number of subsequent iterations of play within the haystack, where the mice there are paired at random. The number of individuals within the haystack choosing each strategy then grows in an endogenous fashion, as does the overall size of the group. Every so often, once a year, say, the haystacks are removed, and the mice are pooled into a single large population. Now pairs of mice are selected at random from the overall population to recolonize the next set of haystacks, and excess mice die.

To give an example, consider the following version of the prisoner’s dilemma—

	<i>C</i>	<i>D</i>
<i>C</i>	2, 2	0, 4
<i>D</i>	4, 0	1, 1

---

<sup>2</sup>See also Robson and Samuelson, 2011.

As another simplification, there are a “large” number of haystacks and therefore a large number of individuals. Suppose that the initial fraction of  $C$ 's in the population is  $f \in [0, 1]$ . Hence the fraction of haystacks that are colonized by 2  $C$ 's is  $f^2$ ; the fraction that are colonized by 2  $D$ 's is  $(1 - f)^2$ ; and the fraction that have one of each is  $2f(1 - f)$ .

There are  $T$  rounds of play within each haystack. In this model, the number of rounds of play within each haystack,  $T$ , will reflect the intensity of group selection relative to individual selection—the larger is  $T$ , the more pronounced will be the advantage of groups with only cooperators relative to all other groups that contain defectors.

It follows that each pair of  $C$ 's gives rise to  $2^{T+1}$  descendants, who are also  $C$ 's. Each pair of  $D$ 's gives rise to just 2  $D$ 's. Each pair of one  $C$  and one  $D$  gives rise to 4  $D$ 's. Choosing the zero for the payoff of the  $C$  against the  $D$  clearly simplifies matters. Furthermore, the payoffs chosen ensure that it always possible to pair all individuals in the haystack.

The new fraction of  $C$ 's is, say,

$$f' = \frac{2^{T+1}f^2}{2^{T+1}f^2 + 8f(1 - f) + 2(1 - f)^2}$$

So it follows that  $f' > f$  if and only if

$$S(f) = 2(1 - f)(3f + 1 - 2^T f) < 0$$

If  $T = 1$ ,  $f' < f$  iff

$$2(1 - f)(1 + f) > 0, \text{ which holds for all } f < 1.$$

That is, in this case, the  $D$ 's will always increase, and  $f \rightarrow 0$ . There is one evolutionarily stable strategy, defection,  $D$ , which is globally stable. This is not surprising, since with  $T = 1$ , we simply have an elaborate description of the usual prisoner's dilemma. Pairs are broken up immediately so that there is no opportunity to exploit the total payoffs for the haystack/group that arise from two initial  $C$ 's.

If  $T = 2$ ,

$$S(f) = 2(1 - f)^2 > 0, \text{ which holds for all } f < 1.$$

Again,  $f \rightarrow 0$  and the  $D$ 's take over the population, for these particular payoffs. There remains one evolutionarily stable strategy, which is globally stable.

If  $T \geq 3$ , however,

$$S(f) = 2(1 - f)(1 - (2^T - 3)f) > 0 \text{ iff } f < \frac{1}{2^T - 3} < 1$$

That is, the fraction of cooperators falls if  $f < \frac{1}{2^T - 3} < 1$ , but increases if  $f > \frac{1}{2^T - 3} > 0$ . That is, in general, there are two evolutionarily stable strategies, one at  $f = 0$  and the other at  $f = 1$ . There is an unstable equilibrium at

$f = \frac{1}{2^T - 3}$ . As the number of repetitions within each haystack increases, the basin of attraction of the defect ESS at  $f = 0$  shrinks and that of the cooperate ESS at  $f = 1$  increases.

The predictions of the theory that allows for a variable intensity of group selection are then: There is always an ESS at  $f = 0$ . This is the prediction of the theory with no group selection. If there is enough group selection, so that the groups persist enough, there will be *another* ESS at  $f = 1$ . The basin of attraction of the second ESS increases with the intensity of group selection, thus becoming the more likely outcome. So the theory that allows for group selection predicts everything that is predicted without group selection and then some.

What empirical content is left to the theory? It is true that the only configurations that would be observed in the prisoner’s dilemma are pure strategies. Is this a general result? The following example shows it is not.

Consider the following version of “chicken”—

	$W$	$H$
$W$	0, 0	2, 4
$H$	4, 2	0, 0

As before, there are a large number of haystacks and therefore individuals. Suppose that the initial fraction of  $W$ ’s in the population is  $f \in [0, 1]$ . Hence the fraction of haystacks that are colonized by 2  $W$ ’s is  $f^2$ ; the fraction that are colonized by 2  $H$ ’s is  $(1 - f)^2$ ; and the fraction that have one of each is  $2f(1 - f)$ . There are again  $T$  rounds of play within each haystack.

The only colonizing pairs that survive the first round within a haystack are then  $\{W, H\}$  which produce 2  $W$ ’s and 4  $H$ ’s right away. These 6 individuals are then paired at random to play the next iteration. If both the  $W$ ’s are paired, which has probability  $1/5$ , the 4  $H$ ’s must also be in like pairs, and the subpopulation dies out. Otherwise, there may exactly one mixed pair  $\{W, H\}$  in which case the next round also has 2  $W$ ’s and 4  $H$ ’s. The remaining possibility is that there are exactly two mixed pairs, in which case, the next round has 4  $W$ ’s and 8  $H$ ’s. That is, either the subpopulation dies out or there are twice as many  $H$ ’s as  $W$ ’s. This argument can be repeated so that at the end of any number of iterations  $T$ , there is a positive probability that the entire subpopulation has died, but also a positive probability that it has not, in which case there are twice as many  $H$ ’s as  $W$ ’s. That is, it may be that strict mixtures of strategies can be maintained by group selection.

Altogether, then, adding a variable (and presumably unknown) intensity of group selection to evolutionary theory will generally increase the range of predictions made, incorporating what will appear as a variable degree of altruism. This decreases the falsifiability of the theory. This disadvantage of such an extended theory may not be seen as automatically fatal. If the effects of group selection were known to be small, however, it would be an argument in favor of keeping the standard economic model.

## 6 Conclusion

Wilson makes a stimulating argument that, in the first place, group selection should be taken more seriously in general. In the second place, and of particular interest to economists, he argues that group selection should be taken especially seriously for humans, since cultural evolution is especially important for us. Indeed, he advocates the strong position that individual humans and human society have a relationship analogous to that between the organs of any individual and that individual. Hence, in particular, economists' view of basic human motivations must be extended to include altruism as a central feature.

Wilson's argument is very clear in that the reader is left in no doubt whatever of Wilson's position on the reality and importance of group selection and altruism. On the other hand, Wilson is less persuasive concerning how the facts support his position. It is ultimately an empirical question, after all. The position he advocates may have merit, at least for humans, but it seems unlikely to have a huge immediate impact on economics, in the absence of compelling empirical evidence.

In this review I argue that the full implications of individual selection for economics have yet to be fully appreciated. I illustrate this by deriving the implications of relative (rather than absolute) fitness for attitudes to risk, showing that this should result in a greater aversion to aggregate risk than to idiosyncratic risk. At the same time, incorporating group selection is bound to reduce the empirical content of economics. I illustrate this by considering the predictions that the haystack model would make in two familiar games.

It is reasonable to advocate, then, that work on the biological foundation for economics provided by orthodox individual selection should be completed before contemplating the possibly radical reshaping implied by utilizing group selection.

## REFERENCES

- Bowles, Samuel. 2006. "Group Competition, Reproductive Leveling, and the Evolution of Human Altruism." *Science* 314(8): 1569-1572.
- Bowles, Samuel. 2009. "Conflict: Nature's Midwife." *Nature* 456: 326-327.
- Boyd, Robert, and Peter J Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Cooper W. S., Kaplan R. H. 1982. "Adaptive 'Coin-Flipping': A Decision-Theoretical Examination of Natural Selection for a Random Individual Variation." *Journal of Theoretical Biology* 94: 135-15.
- Curry, Philip. 2001. "Decision Making under Uncertainty and the Evolution of Interdependent Preferences," *Journal of Economic Theory* 98: 357-369.
- Darwin, C. 1871. *The Descent of Man and Selection in Relation to Sex*. London: John Murray.
- Dawkins, R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

- Grafen, Alan. 1991. "Modelling in behavioural ecology." Ch. 1 in *Behavioral Ecology: An Evolutionary Approach*, ed. J. R. Krebs and N. B. Davies. 3rd ed. Oxford: Blackwell Scientific.
- Grafen, Alan. 1999. Formal Darwinism, the individual-as-maximising-agent analogy, and bet-hedging. *Proceedings of the Royal Society Series B* 266: 799–803.
- Maynard Smith, John. 1964. "Group Selection and Kin Selection." *Nature* 201: 1145-1147.
- Mayr, Ernst. 1991. *One Long Argument*. Cambridge, MA: Harvard University Press.
- Popper, Karl. 2002. *The Logic of Scientific Discovery*. London and New York: Routledge.
- Robalino, Nikolaus, and Arthur J. Robson. 2013. "Genes, Culture, and Preferences." *Biological Theory* 8: 151-57.
- Robson, Arthur J. 1996. "A Biological Basis for Expected and Non-Expected Utility." *Journal of Economic Theory* 68: 397-424.
- Robson, Arthur J. 2008. "Group Selection" in *The New Palgrave: A Dictionary of Economics, Second Edition*. London: Palgrave Macmillan.
- Robson, Arthur J. and Larry Samuelson. 2011. "The Evolutionary Foundations of Preferences" In *Handbook of Social Economics*. Edited by Alberto Bisin and Matt Jackson, North-Holland, 221-310.
- Schaffer, Mark E. 1988. "Evolutionarily Stable Strategies for a Finite Population and a Variable Contest Size." *Journal of Theoretical Biology* 132(4): 469–478.
- Smith, Adam. 1761. *Theory of Moral Sentiments*. (2nd ed.) Strand & Edinburgh: A. Millar; A. Kincaid & J. Bell.
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations* (1st ed.) London: W. Strahan.
- Sober, Eliot, and David Sloan Wilson. 1999. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Williams, George C. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton: Princeton University Press.
- Wilson, David Sloan. 2015. *Does Altruism Exist?* New Haven: Yale University Press.
- Wynne-Edwards, V.C. 1962. *Animal Dispersion in Relation to Social Behavior*. Edinburgh: Oliver and Boyd.