



Cite this article: Peters O, Adamou A. 2022

The ergodicity solution of the cooperation puzzle. *Phil. Trans. R. Soc. A* **380**: 20200425. <https://doi.org/10.1098/rsta.2020.0425>

Received: 30 April 2021

Accepted: 20 October 2021

One contribution of 17 to a theme issue 'Emergent phenomena in complex physical and socio-technical systems: from cells to societies'.

Subject Areas:

complexity, mathematical physics

Keywords:

cooperation, ergodicity, evolution

Author for correspondence:

Ole Peters

e-mail: o.peters@lml.org.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5885340>.

The ergodicity solution of the cooperation puzzle

Ole Peters^{1,2} and Alexander Adamou¹

¹London Mathematical Laboratory, 8 Margravine Gardens, London W6 8RH, UK

²Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

OP, 0000-0002-1246-8570; AA, 0000-0002-4188-7045

When two entities cooperate by sharing resources, one relinquishes something of value to the other. This apparent altruism is frequently observed in nature. Why? Classical treatments assume circumstances where combining resources creates an immediate benefit, e.g. through complementarity or thresholds. Here we ask whether cooperation is predictable without such circumstances. We study a model in which resources self-multiply with fluctuations, a null model of a range of phenomena from viral spread to financial investment. Two fundamental growth rates exist: the ensemble-average growth rate, achieved by the average resources of a large population; and the time-average growth rate, achieved by individual resources over a long time. As a consequence of non-ergodicity, the latter is lower than the former by a term which depends on fluctuation size. Repeated pooling and sharing of resources reduces the effective size of fluctuations and increases the time-average growth rate, which approaches the ensemble-average growth rate in the many-cooperator limit. Therefore, cooperation is advantageous in our model for the simple reason that those who do it grow faster than those who do not. We offer this as a candidate explanation for observed cooperation in rudimentary environments, and as a behavioural baseline for cooperation more generally.

This article is part of the theme issue 'Emergent phenomena in complex physical and socio-technical systems: from cells to societies'.

1. Introduction

They give that they may live, for to withhold is to perish [1].

Living beings exist not as minimal self-reproducing units but as cells, organisms, families, institutions, nations and so on. Cooperation, which we model as sharing resources, is ubiquitous in nature and society.

This ubiquity is puzzling because sharing seems *prima facie* to require the better-off member of a cooperating pair to relinquish something of value to the worse-off member, with nothing in return. If naked altruism is an unsatisfactory explanation of evolved behaviour, then we must expose the advantage derived by the better-off entity in such an arrangement.

Classical explanations involve two ideas. The first is that a net benefit arises when two entities cooperate. Specifically, the gain of the recipient—often expressed in terms of ‘fitness’—exceeds the cost to the donor. The second is that, over time, some of the net benefit finds its way back to the donor. This can happen through reciprocity, where past donors become future recipients, or through relatedness, where the recipient carries genetic material that the donor wants to propagate. Nowak [2] offers a comprehensive account of this approach, describing cooperative structures that are possible when such benefits exist. Aktipis [3] explores the regulatory mechanisms that facilitate such structures at large scales and over long times.

A net benefit typically requires complementarity in the shared resources, or a resource threshold below which a desired outcome is unachievable. Complementarity can arise through differences in resource types, such as the knowledge and materials required to send a human to the moon, or through splitting tasks into complementary parts, such as one hunter chasing prey towards the other. Thresholds are common in human economic activity, e.g. business partners combining their resources to purchase a productive asset, such as manufacturing equipment, which each could not afford alone. In all cases, attendant circumstances are such that the whole is deterministically ‘more than the sum of its parts’.

Here we ask whether cooperation can arise without such effects. We study this question theoretically in a model population whose members’ resources follow noisy multiplicative growth, i.e. periodic multiplication by a random factor. For tractability, we use geometric Brownian motion (GBM), a continuous-time stochastic process in which resource multipliers are normally distributed. GBM is the limiting case of a broad class of multiplicative growth processes. Its universality makes it suitable for modelling a wide range of phenomena, from the spread of a virus to the compounding of investments. The message of this study is general and we refer abstractly to ‘populations’, ‘entities’ and ‘resources’, allowing readers to decide what interests them specifically. Importantly, resources in our model are identical in all respects other than their random fluctuations. This allows us to study cooperation under simple conditions, where the effects invoked in classical explanations do not exist.

There are two fundamental growth rates in noisy multiplicative growth. The *ensemble-average growth rate* is that achieved by the population average in the large-population limit. It is independent of fluctuations. The *time-average growth rate* is that achieved by a single entity in the long-time limit. It is lower than the ensemble-average growth rate by a fluctuation-dependent term. The difference between these growth rates is a manifestation of non-ergodicity; see [4] and references therein. We find that repeated pooling and sharing reduces the net effect of fluctuations, thereby increasing the time-average growth rate of each cooperator’s resources, which approaches the ensemble-average growth rate as the number of cooperators increases. Therefore, cooperation in our model is advantageous for the simple reason that those who do it outgrow those who do not.

We suggest two major implications of this theoretical finding. Firstly, it provides a candidate explanation for observed cooperation in real settings too simple for classical explanations to hold. One tantalizing example is the transition from unicellular to multicellular life. Classical explanations usually involve the emergence of some new function, e.g. the ability of a

multicellular organism to swim up a nutrient gradient [5,6]. However, a new function requires a degree of complexity absent in early unicellular life. A credible theory of evolution must explain not only the rich tapestry of cooperative structure we see today but also early cooperative steps—such as from single cell to cell pair—in rudimentary environments. Similar questions arise in the formation of early human societies and in human activities prior to advanced technologies, such as land management before the agricultural revolution [7].

Secondly, our finding establishes a leading-order result, or behavioural baseline, from which deviations can be studied. If, as we suggest, cooperation is beneficial by default in a model with minimal assumptions, then special explanations are required when it is not observed in nature. Moreover, by showing that cooperation has a quantifiable benefit to individuals who engage in it, this result counsels against overly narrow definitions of self-interest.

Our work relates closely to the biological literature on geometric mean fitness, on which the negative effects of fluctuations have long been recognized [8–11]. We speak of growth rates and not fitness, since they are physically measurable and unequivocally defined, whereas fitness has multiple definitions [11]. Among entities whose resources undergo noisy multiplicative growth, those with the highest time-average growth rates will dominate their environment over time.

Also related is the literature on bet-hedging, in which sharing or diversification mechanisms are employed to reduce the variance of resource yields, often at the cost of reducing their mean. Phenomena such as land scattering by peasant farmers [7], food sharing by foragers [12] and, more recently, helping behaviours of birds [13] are explained as variance-reduction strategies, with the aim of reducing the frequency of catastrophic shortfalls and prolonging survival. Uitdehaag [14] explores a model similar to ours, in which multiplicative growers share alternately their surpluses under a no-net-donor constraint. He finds that fluctuation-induced reductions in proliferative success are best countered by sharing between entities with anticorrelated fluctuations, proposing this as an explanation of mutualism between specialists. Starrfelt & Kokko [15] formalize bet-hedging as a trade-off between means, variances and correlations of resource increments. Schreiber [16] relates this to the probability of bet-hedger genotypes dominating a population over time, while Kennedy *et al.* [17] express it as a special case of a general model of cooperation which allows for relatedness between entities [18].

We present our work as follows. In §2, we introduce our model of noisy multiplicative growth and review its properties. In §3, we describe our cooperation protocol in which entities grow, pool and share their resources repeatedly. In §4, we find that cooperation increases the time-average growth rate of resources of similar entities and we discuss whether this constitutes a type of group selection. In §5, we discuss generalizations of our model to dissimilar entities and correlated fluctuations. Finally, we summarize our findings in §6.

2. Noisy multiplicative growth

Here we introduce our underlying model of resource growth in the absence of cooperation, to which we add a resource-sharing protocol in §3.

Let $x_i(t)$ be the resources of entity i at time t . We assume each $x_i(t)$ follows noisy multiplicative growth, meaning that it is multiplied periodically by realizations of a random variable. While we speak abstractly of the resources of an entity, our analysis is agnostic to replacement of ‘resources’ by ‘biomass’, ‘food’, ‘wealth’ and so on, and of ‘entity’ by ‘cell’, ‘organism’, ‘colony’, ‘person’, ‘tribe’ and so on. Cooperation occurs in many domains and at many scales.

Our specific model is GBM, where the change in resources over a short time step is a normally distributed random multiple of the existing resources. More formally, $x_i(t)$ follows the Itô stochastic differential equation,

$$dx_i(t) = x_i(t) (\mu dt + \sigma dW_i(t)), \quad (2.1)$$

where μ is the drift and σ is the volatility. The $dW_i(t)$ are independent and identically distributed increments of the Wiener process which are normal variates with zero mean and variance dt .

The realizations $dW_i(t)$ in equation (2.1) are indexed by both i and t . This is because we anticipate different fluctuations for different entities and at different times, for example due to environmental conditions that vary over space and time. This spatio-temporal distinction is important in the biological literature, where within-generation and between-generation variances of outcomes are studied [15,16]. Furthermore, we assume the $dW_i(t)$ are independent. We explore the effects of correlated fluctuations, i.e. of entities experiencing consistently similar or dissimilar conditions, in §5.

GBM is a universal model because it is an attractor for more complex models of multiplicative growth. Since its logarithmic increments are normally distributed, any stochastic process to whose logarithmic increments the central limit theorem applies will be distributed as a GBM in the many-increment limit. Formal treatments can be found in [19,20]. The key point is that there exists a broad class of multiplicative processes which, over time scales much longer than a single time step, are well approximated by GBM.

We note that GBM is a model of unconstrained self-reproduction. Growth that is limited by resource or space constraints or by predation would be poorly described by equation (2.1). The water lily of the famous riddle, told in [21], stops growing exponentially once it covers the pond.

It is an established property of GBM, which we quote without derivation, that its solution follows a time-dependent lognormal distribution,

$$\ln x_i(t) \sim \mathcal{N} \left(\ln x_i(0) + \left(\mu - \frac{\sigma^2}{2} \right) t, \sigma^2 t \right), \quad (2.2)$$

where $\mathcal{N}(m, s^2)$ denotes the normal distribution with mean m and variance s^2 . Therefore, over a time period T , each entity's resources experience a random exponential growth rate,

$$g(x_i, T) \equiv \frac{1}{T} \ln \left(\frac{x_i(T)}{x_i(0)} \right), \quad (2.3)$$

which follows a normal distribution,

$$g(x_i, T) \sim \mathcal{N} \left(\mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{T} \right). \quad (2.4)$$

Imagine starting many cell cultures in separate Petri dishes and watching their biomasses evolve according to equation (2.1) for time T . Assume the dishes are large enough and T is short enough that growth does not slow for want of agar. Equation (2.4) says that the observed growth rates would be normally distributed around $\mu - \sigma^2/2$, with time-decaying variance σ^2/T .

The negative term $-\sigma^2/2$ may be surprising, given that the multiplicative changes in equation (2.1) are distributed symmetrically about μdt . It is, however, a well-known property, usually derived using Itô calculus as in [22]. In simple terms, it reflects the fact that the product of the symmetric perturbations of a number is less than its square, e.g. $(1 + \epsilon)(1 - \epsilon) = 1 - \epsilon^2 < 1$.

The ensemble average (or expectation value) of resources is defined as

$$\langle x(t) \rangle \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i(t). \quad (2.5)$$

Its evolution is computed by applying this operation to equation (2.1) to get

$$d\langle x(t) \rangle = \mu \langle x(t) \rangle dt, \quad (2.6)$$

since $\langle dW_i(t) \rangle = 0$. With initial condition $x_i(0) = 1$ for all i , this describes exponential growth, $\langle x(t) \rangle = \exp(\mu t)$, at a rate equal to the drift, μ . We will call this the *ensemble-average growth rate* and denote it by

$$g(\langle x \rangle) = \mu. \quad (2.7)$$

The physical interpretation of this quantity is worth making explicit: it is the growth rate of the average resources of a population in the large-population limit.

One might guess that the growth rate observed in an individual trajectory will converge to equation (2.7) over time, but this would be a common conceptual error. Instead, the non-ergodicity of equation (2.1) manifests itself such that the growth rate observed in a single trajectory converges to a different value, called the *time-average growth rate*. This is the $T \rightarrow \infty$ limit of equation (2.4), in which the variance decays almost surely to zero to leave

$$\bar{g}(x_i) \equiv \lim_{T \rightarrow \infty} g(x_i, T) = \mu - \frac{\sigma^2}{2}. \quad (2.8)$$

We see in nature and society what has survived. In our model, the entity with the highest time-average growth rate will, regardless of its ensemble-average growth rate, come to dominate the environment's resources in the long-time limit. The ratio of its resources to those of other entities will grow exponentially. Strategies which increase $\bar{g}(x_i)$, regardless of their effect on $g(\langle x \rangle)$, will confer an evolutionary advantage on their adherents. We find that cooperation, specifically the repeated pooling and sharing of resources, is one such strategy.

3. Cooperation protocol

Having established the properties of equation (2.1), we introduce our model of cooperation. We start with a population of N non-cooperating entities, whose resources follow GBM with identical drift and volatility, i.e. equation (2.1) with $i = 1, \dots, N$.

For simplicity, we consider a discrete version of equation (2.1). The non-cooperators' resources grow over a finite time step, Δt , according to

$$\Delta x_i(t) = x_i(t) \left(\mu \Delta t + \sigma \xi_i(t) \sqrt{\Delta t} \right) \quad (3.1)$$

and

$$x_i(t + \Delta t) = x_i(t) + \Delta x_i(t), \quad (3.2)$$

where $\xi_i(t)$ are independent standard normal variates, $\xi_i(t) \sim \mathcal{N}(0, 1)$.

The cooperation mechanism, summarized pictorially for $N=2$ in figure 1, is as follows. Previously independent entities with resources $x_i(t)$ start to cooperate. This might happen through a genetic mutation which forces entities to share resources, as in the context of primitive life, or through more complex regulatory mechanisms, as in animal and human interactions. We label the resources of the cooperating entities $y_i(t)$ to distinguish them from the corresponding non-cooperators. We assume equal sharing of resources, $y_i(t) = y^{(N)}(t)$ for all i , where $y^{(N)}(t)$ denotes the per-entity resources for N equal cooperators.

In the discrete-time picture, each time step involves two phases. First, there is a growth phase, analogous to equation (3.1), in which each cooperator increases its resources by

$$\Delta y_i(t) = y^{(N)}(t) \left(\mu \Delta t + \sigma \xi_i(t) \sqrt{\Delta t} \right). \quad (3.3)$$

This is followed by a cooperation phase, replacing equation (3.2), in which resources are pooled and shared equally among the cooperators:

$$y^{(N)}(t + \Delta t) = \frac{1}{N} \sum_{i=1}^N (y^{(N)}(t) + \Delta y_i(t)) = y^{(N)}(t) + \frac{1}{N} \sum_{i=1}^N \Delta y_i(t). \quad (3.4)$$

Equation (3.4) is equivalent to equal sharing of the sum of the individual fluctuations,

$$\Delta y^{(N)}(t) = \frac{1}{N} \sum_{i=1}^N \Delta y_i(t). \quad (3.5)$$

Our model of cooperation is necessarily technical. A lay version, presented as an interactive fable about farmers who boost their grain yields by repeated pooling and sharing, is available at www.farmersfable.org.

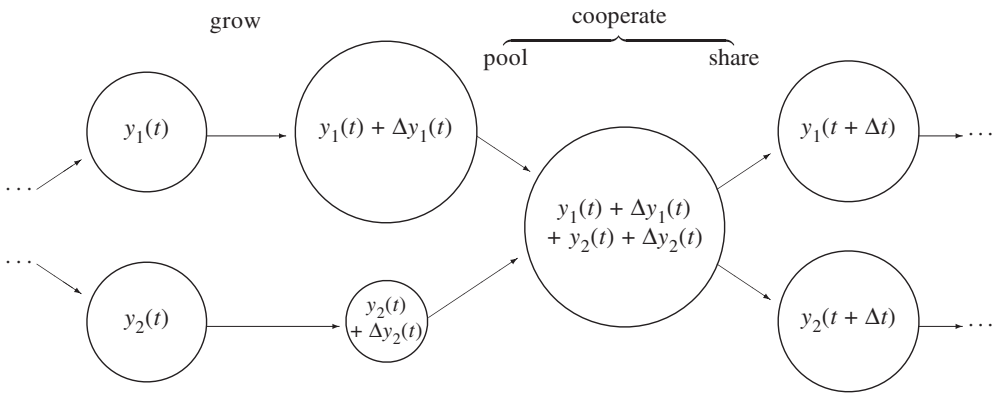


Figure 1. Cooperation dynamics. Two cooperators start each time step with equal resources. Then they grow independently according to equation (3.3). Then they cooperate by pooling resources and sharing them equally, according to equation (3.4). Then the next time step begins.

Cooperation has no direct cost in this protocol. In reality, sharing often requires a coordinating mechanism. For example, large organisms have circulatory systems to disseminate nutrients, and human societies have administrative systems to redistribute resources. Such mechanisms may have costs that make cooperation disadvantageous [3]. Equally, we ascribe no direct benefit to cooperation. Costs and benefits emerge as the effects of cooperation on time-average growth rates.

Nor do we consider the cheating problem or the walk-away option [3], although we will discuss later when breaking away will be beneficial. For now, our cooperators are unable to break their pact.

Substituting equation (3.3) into equation (3.5) yields the dynamic followed by the resources of each cooperator

$$\Delta y^{(N)}(t) = y^{(N)}(t) \left(\mu \Delta t + \frac{\sigma}{N} \sum_{i=1}^N \xi_i(t) \sqrt{\Delta t} \right). \quad (3.6)$$

The simplification of the final term leads to the key insight of this paper. Sums of independent normal variates are normal, so we can define a single standard normal variate,

$$\xi^{(N)}(t) \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^N \xi_i(t) \sim \mathcal{N}(0, 1), \quad (3.7)$$

and rewrite equation (3.6) as

$$\Delta y^{(N)}(t) = y^{(N)}(t) \left(\mu \Delta t + \frac{\sigma}{\sqrt{N}} \xi^{(N)}(t) \sqrt{\Delta t} \right). \quad (3.8)$$

Thus, the net effect of N individual fluctuations pooled and shared is a single equivalent fluctuation, whose amplitude is $1/\sqrt{N}$ times the amplitude of the individual fluctuations. Substituting into equation (3.6) and letting the time increment $\Delta t \rightarrow 0$, we recover a stochastic differential equation of the same form as equation (2.1) and with the same drift μ , but with the volatility reduced from σ to $\sigma/\sqrt{N}$,

$$dy^{(N)}(t) = y^{(N)}(t) \left(\mu dt + \frac{\sigma}{\sqrt{N}} dW^{(N)}(t) \right). \quad (3.9)$$

The effect of this volatility reduction on the time-average growth rate of resources provides our answer to the cooperation puzzle, as we shall now show.

4. Ergodicity solution of the cooperation puzzle

The ensemble-average growth rates of $x_i(t)$ under equation (2.1) and $y^{(N)}(t)$ under equation (3.9) are identical: μ . From this perspective, there is no incentive to cooperate. Moreover, an entity that realizes a large positive resource change during a growth phase, equation (3.3), could keep its fortuitous gain by breaking the cooperative pact, without altering its ensemble-average growth rate. Thus, analysing the growth of the ensemble average of resources gives no reason for cooperation to arise and, if it does arise, a good reason for it to end. From this perspective, cooperation looks fragile at best and its frequent observation in nature seems puzzling.

The solution of the puzzle requires a change in perspective. The ensemble-average growth rate is uninformative of the growth of a single entity (or, indeed, any finite cooperative) over time. From the perspective of an individual, it is the growth rate achieved by averaging over infinitely many parallel realizations of its resources—an irrelevant fiction. It is more realistic to assume that an individual cares about its time-average growth rate, which better approximates what it will actually achieve in one resource trajectory over time.

We know from equation (2.8) that the time-average growth rate for non-cooperating entities is $\bar{g}(x_i) = \mu - \sigma^2/2$. Under cooperative dynamics, equation (3.9), the volatility decreases from σ to $\sigma/\sqrt{N}$, which when substituted into equation (2.8) yields a higher time-average growth rate,

$$\bar{g}(y^{(N)}) = \mu - \frac{\sigma^2}{2N}. \quad (4.1)$$

So, in the presence of fluctuations, $\sigma > 0$, cooperators grow faster than non-cooperators in the long run. The growth rate premium increases with the number of cooperators as $1 - 1/N$,

$$\bar{g}(y^{(N)}) - \bar{g}(x_i) = \frac{\sigma^2}{2} \left(1 - \frac{1}{N}\right), \quad (4.2)$$

implying that larger cooperatives are favoured over smaller ones. This premium increases most rapidly when N is small—from single entities, to pairs, to triplets. As the cooperative expands, the benefit gained by adding each new member diminishes [12].

In our model, cooperators will eventually dominate the environment and cooperation will become ubiquitous. The effect is illustrated in figure 2 by direct simulation of equations (3.1), (3.2) and (3.6). It constitutes our main result. We see that over a long enough time, the resources of each member of a cooperating pair (blue) grow faster than the resources of the corresponding non-cooperators (green). They also grow faster than the average resources of the non-cooperators (black), showing that cooperation and averaging are not equivalent operations.

In equation (4.1), we see that the time-average growth rate of a cooperator approaches the ensemble-average growth rate of a non-cooperator as the number of cooperators grows large,

$$\lim_{N \rightarrow \infty} \bar{g}(y^{(N)}) = \mu. \quad (4.3)$$

The pink dashed line in figure 2 plots resource growth at the ensemble-average growth rate. This is relevant to members of large cooperatives, equation (4.3), but bears little resemblance to that of the resource growth of non-cooperators (green) and members of small cooperatives (blue). This is the essence of the ergodicity debate: features emerging from fluctuations when averaging over a statistical ensemble do not, in general, also emerge in individual systems over time.

(a) The group selection debate

One explanation of observed cooperative behaviour—known as group or multi-level selection—has led to controversy among evolutionary biologists; see [23] and references therein. While it is unclear whether the debate is semantic or substantive [24], we nevertheless offer a brief summary of it and its relation to our work.

The dominant paradigm in evolutionary biology, originating in [25], is that natural selection acts at the level of the replicating unit, such as an organism or its genotype. We call this

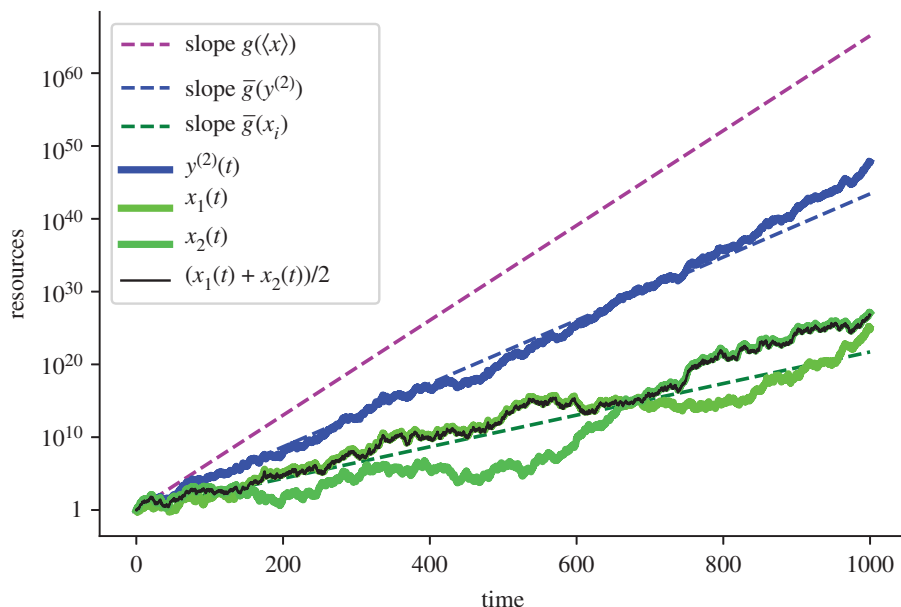


Figure 2. Typical trajectories for two non-cooperating (green) entities and for the corresponding cooperating entities (blue) on a logarithmic resource scale. Over time t , the volatility reduction achieved by the cooperators leads to faster growth. The thin black line shows the average resources of the non-cooperators, which is far inferior to those of the cooperators. In a literal mathematical sense, the whole, $y_1(t) + y_2(t)$, is more than the sum of its parts, $x_1(t) + x_2(t)$. Resource growth at the ensemble-average growth rate (pink) is approached in the limit of infinitely many cooperators. Parameters: $\mu = 0.15$ per time unit, $\sigma^2 = 0.2$ per time unit. The discrete time step Δt was set to 0.1 time units. The code to generate the figure is in the electronic supplementary material, `cooperate.py`. (Online version in colour.)

individual selection to distinguish it from group selection. Put simply, a trait which increases the reproductive success of individuals over those with other traits will tend to increase its prevalence in a population over time.

The consensus view is that individual selection explains observed cooperation [26]. The premise is that all cooperative behaviour is beneficial to the individual engaging in it. If such behaviour is interpreted as *prima facie* disadvantageous, this is simply a mistake arising from an over-narrow view of the behaviour's consequences or the individual's interests. Reciprocity and relatedness are examples where apparently costly aid is actually self-interested: either because it leads to receipt of aid in the future, or because it benefits the genotype of the aid-giver. Indeed, the cooperation puzzle presented here may be viewed as an artefact of ignoring the wider circumstances of a resource-sharing act, namely that it belongs to a sequence of similar acts.

Group selection is an alternative paradigm in which natural selection occurs at group level [27–29]. Behaviour which is apparently costly to the individual is rendered unproblematic, since it is the group to which selection pressure applies. Thus, individual traits can evolve which are selectively good for the group but selectively bad for individuals within the group. This approach is controversial *inter alia* because groups are not considered replicating units and because individual selection is already deemed sufficient to explain cooperation. We offer no new analysis of this controversy, which is beyond the scope of this paper.

However, we suggest that our model of cooperation sits naturally in the individual-selection paradigm. Individuals pool and share resources because, over a temporal sequence of interactions, it is individually advantageous for them to do so. Behavioural traits are good for the group precisely because they are good for every member of the group. Group selection is not needed because the interests of the group and its members are fully aligned.

This raises the question of whether group and individual interests could become misaligned in our model. For example, could invasion of a resource-sharing cooperative by exploitative non-cooperators create a situation where cooperators engage in personally costly behaviour to benefit the group or a subset thereof?

Leaving aside the question of whether such a mixed group would be sustainable, which would depend on the exploitation mechanism, it is clear that breaking the cooperative pact once creates a short-term advantage for the pact-breaker. However, provided regulation exists to prohibit repeated pact-breaking, this will not translate into a long-term advantage. One simple way in which cooperators could protect themselves from repeated exploitation would be to recognize past exploiters and refuse to interact with them in future—an ‘ignore-for-tat’ policy.

The penalty for non-cooperation would be ejection from the cooperative, reduction of the time-average growth rate from $\mu - \sigma^2/(2N)$ to $\mu - \sigma^2/2$ and the eventual exponential irrelevance this implies. Detection and punishment need not be perfect. Even if the penalty were exacted with non-zero probability at each betrayal, it would be only a matter of time before the traitor was expelled from the group and consigned to lag behind. However, since our main aim is to study *why* resource-sharing has an evolutionary advantage, rather than *how* it is enforced, we refrain from further development of this point.

5. Generalizations

(a) Idiosyncratic entities

Real cooperatives have members of differing abilities as well as differing fortunes. The latter we model already as different realizations of the Wiener increments in equation (2.1). The former we can treat by generalizing equation (2.1) so that the entities have idiosyncratic drifts and volatilities,

$$dx_i(t) = x_i(t)(\mu_i dt + \sigma_i dW_i(t)) \quad (5.1)$$

for $i = 1, \dots, N$. The time-average growth rates, $\bar{g}(x_i) = \mu_i - \sigma_i^2/2$, are now idiosyncratic. Some entities will, if left to their own devices, grow their resources faster than others. This raises questions. Does it benefit leaders to share with laggards? When should a non-cooperator join a cooperating group? When should the group allow it?

Repeating the analysis of growth, pooling and sharing yields a modified dynamic

$$\begin{aligned} \Delta y^{(N)}(t) &= y^{(N)}(t) \left(\frac{1}{N} \sum_{i=1}^N \mu_i \Delta t + \frac{1}{N} \sum_{i=1}^N \sigma_i \xi_i(t) \sqrt{\Delta t} \right) \\ &= y^{(N)}(t) \left(\mu^{(N)} \Delta t + \sigma^{(N)}(t) \xi^{(N)} \sqrt{\Delta t} \right), \end{aligned} \quad (5.2)$$

where $\xi^{(N)}(t)$ is a standard normal variate, as before, and

$$\mu^{(N)} \equiv \frac{1}{N} \sum_{i=1}^N \mu_i, \quad \sigma^{(N)} \equiv \frac{1}{N} \sqrt{\sum_{i=1}^N \sigma_i^2} \quad (5.3)$$

are the effective drift and volatility parameters. Therefore, the resources of the cooperators evolve according to

$$dy^{(N)}(t) = y^{(N)}(t)(\mu^{(N)} dt + \sigma^{(N)} dW^{(N)}(t)), \quad (5.4)$$

with time-average growth rate

$$\bar{g}(y^{(N)}) = \mu^{(N)} - \frac{(\sigma^{(N)})^2}{2} = \frac{1}{N} \sum_{i=1}^N \left(\mu_i - \frac{\sigma_i^2}{2N} \right). \quad (5.5)$$

This happens to be the sample mean of the time-average growth rates each entity would achieve in a cooperative of N like entities; cf. equation (4.1).

We can now answer the questions. It benefits entity j to be a member of the cooperative if $\bar{g}(y^{(N)}) > \bar{g}(x_j)$, i.e. if

$$\frac{1}{N} \sum_{i=1}^N \left(\mu_i - \frac{\sigma_i^2}{2N} \right) > \mu_j - \frac{\sigma_j^2}{2}. \quad (5.6)$$

Similarly, the cooperative benefits by having j as a member if

$$\frac{1}{N} \sum_{i=1}^N \left(\mu_i - \frac{\sigma_i^2}{2N} \right) > \frac{1}{N-1} \sum_{\substack{i=1 \\ i \neq j}}^N \left(\mu_i - \frac{\sigma_i^2}{2(N-1)} \right). \quad (5.7)$$

In reality, we may interpret this as saying that, even in simple set-ups, a better-skilled individual motivated by nothing but greed can still do better as part of society.

(b) Correlated fluctuations

A second generalization concerns correlations. Fluctuations experienced by different entities are uncorrelated in our model. The $dW_i(t)$ in equation (2.1) and, consequently, the $\xi_i(t)$ in equation (3.1) onwards are independent random variables. In reality, cooperators are often spatially localized or socially connected, and they experience similar environmental or economic conditions at a given time. By allowing correlations in fluctuations across entities, i.e. correlated $\xi_i(t)$ and $\xi_j(t)$, our model can be adapted to describe such situations.

Suppose the $\xi_i(t) \sim \mathcal{N}(0, 1)$ realized in a given time step, from t to $t + \Delta t$, are jointly normal and cross-correlated such that $\langle \xi_i(t) \xi_j(t) \rangle = \rho_{ij}$. Assume for simplicity that

$$\rho_{ij} = \begin{cases} 1, & i = j, \\ \rho, & i \neq j, \end{cases} \quad (5.8)$$

so that the fluctuations for all pairs of different entities have the same covariance, ρ , where $-1 \leq \rho \leq 1$. The more general case of a covariance matrix with unequal off-diagonal elements is also tractable, but adds complexity without illumination.

The presence of cross-correlations alters the evaluation of the sum of the normal variates in equation (3.6). We have now

$$\sum_{i=1}^N \xi_i(t) \sim \mathcal{N}(0, N + N(N-1)\rho). \quad (5.9)$$

Positive variance requires ρ to be confined to $-1/(N-1) \leq \rho \leq 1$ (meaning that perfect anticorrelation can exist only for $N=2$ and that anticorrelation is impossible as $N \rightarrow \infty$). Equation (5.9) suggests defining, analogously to equation (3.7), a standard normal variate

$$\xi^{(N)}(t) \equiv \frac{1}{\sqrt{N + N(N-1)\rho}} \sum_{i=1}^N \xi_i(t) \sim \mathcal{N}(0, 1), \quad (5.10)$$

such that the change in $y^{(N)}(t)$ can be written as

$$\Delta y^{(N)}(t) = y^{(N)}(t) \left(\mu \Delta t + \sigma \sqrt{\frac{1 + (N-1)\rho}{N}} \xi^{(N)}(t) \sqrt{\Delta t} \right), \quad (5.11)$$

analogous to equation (3.8) in the uncorrelated case.

Without correlations, cooperation reduces the amplitude of fluctuations from σ to $\sigma/\sqrt{N}$. With them, the fluctuation amplitude becomes

$$\sigma_{\text{corr}} \equiv \sigma \sqrt{\frac{1 + (N-1)\rho}{N}}. \quad (5.12)$$

McCloskey [7] obtains the same expression for the variance of the annual yield of peasant farmers with scattered plots, as does Winterhalder [12] for teams of foragers sharing food.

The variation of σ_{corr} with ρ and N delineates the main features of this model. Firstly, as a consistency check, we note that $\sigma_{\text{corr}} \rightarrow \sigma/\sqrt{N}$ as $\rho \rightarrow 0$ for fixed N , recovering the uncorrelated result in the appropriate limit.

For all $N > 1$, we have $0 \leq \sigma_{\text{corr}} \leq \sigma$, with $\sigma_{\text{corr}} = \sigma$ if and only if $\rho = 1$. In other words, provided fluctuations are not perfectly correlated, a cooperation benefit always exists. This makes intuitive sense. With perfect correlation, all the ξ_i are identical and sharing achieves nothing. The cooperative is equivalent to a giant individual following a single trajectory of equation (2.1). As soon as some variation is introduced between the fluctuations of the entities, the cooperation mechanism can begin to mitigate the negative effects of fluctuations on resource growth.

Furthermore, in the $N \rightarrow \infty$ limit we have $\sigma_{\text{corr}} \rightarrow \sigma\sqrt{\rho}$ with $\rho \geq 0$. The maximum time-average growth rate achievable by adding cooperators is, therefore,

$$\lim_{N \rightarrow \infty} \bar{g}(y^{(N)}) = \mu - \frac{\sigma^2 \rho}{2}; \quad (5.13)$$

cf. equation (4.3). This cannot exceed μ and decreases as ρ increases. Again, this is consistent with intuition: as fluctuations become more correlated, the variation between them diminishes and the scope for beneficial cooperation narrows.

In our set-up, cooperation is enhanced by diversity in individual outcomes. Uitdehaag [14] illustrates the benefit of specialization in mutualistic societies with an extreme case of two entities, $N = 2$, with perfect anticorrelation, $\rho = -1$. Here, $\sigma_{\text{corr}} = 0$ and the time-average growth rate, $\mu - \sigma_{\text{corr}}^2/2$, becomes the ensemble-average growth rate, μ , with no further cooperation required. Indeed, maintaining diversity in outcomes is the explanation for land scattering favoured by McCloskey [7], citing first-hand accounts of reasoning given by farmers engaging in the practice. Similarly, where cooperation imposes structures that lead to a loss in diversity, the cooperation benefit is diminished and can be completely eliminated.

(c) Other generalizations

Many other generalizations are possible. Berman *et al.* [30] analyse partial cooperation, in which entities pool and share a fraction of their resources, resembling taxation and redistribution in human societies. Stojkoski *et al.* [31] consider cooperation on a network, where entities pool and share resources only with those to whom they are connected. Yaari & Solomon [32] find an evolutionary advantage of cooperation in a different multiplicative process, with only two possible outcomes (win or loss) at each step, and we anticipate that similar results will hold for more general resource processes, such as those described by the present authors in [33].

6. Discussion

The cooperation puzzle refers to the common observation of spontaneous cooperative behaviour which, at first sight, appears unattractive to one cooperator. Classical treatments resolve it under circumstances where cooperation creates an immediate benefit, for instance where resources are complementary or where thresholds exist that an entity cannot surpass alone.

We have presented a simple and universal model of resource growth and sharing where cooperation proves beneficial in the absence of such circumstances. The time-average growth rate of resources is reduced by fluctuations. Pooling and sharing, by decreasing the effective amplitude of fluctuations, increases the time-average growth rate. Thus, cooperation is advantageous in our model because those who do it grow faster than those who do not. The simplicity and generality of our model make this a candidate explanation for the existence of cooperation in a wide range of real settings, especially in simple environments where classical explanations are unlikely to hold.

This minimalism allows us to ask which of cooperation and non-cooperation is the behavioural baseline and, therefore, which should be considered puzzling when observed. Following our analysis, we expect cooperation to be attractive when resources are, broadly speaking, self-reproductive—a class of situations encompassing the biomass of bacteria and the wealth of

nations. In such situations, the absence of cooperation requires special explanations, such as coordination costs, diversity suppression (e.g. ‘groupthink’) or skill differences among would-be cooperators that discourage the most skilled from participating.

The impact of fluctuation reduction on long-time growth suggests that risk management has a rarely recognized significance. Good risk management does not merely reduce the size of up- and down-swings. It also improves long-time performance. This tangible economic benefit is frequently overlooked. The coordination costs and inefficiencies of insurance, pension and taxation systems are often discussed, without passing mention of this fundamental benefit that, we suppose, had much to do with their emergence in the first place.

Cooperation has no apparent benefit when judged by its effect on the ensemble-average growth rate of resources. Only the time-average growth rate reveals the benefit we have described. This work forms part of a research programme known as ergodicity economics; see e.g. [4]. This programme uses the ergodicity problem as a lens through which it redevelops the dominant economic formalism. The development of ergodic theory in the twentieth century provided the concepts needed to analyse stochastic growth processes like equation (2.1). Such analysis suggests that our natural tendency to cooperate—expressed in our gut feeling and moral sentiment—is in harmony with a careful formal analysis.

Data accessibility. The data are provided in the electronic supplementary material [34]. Data displayed in figure 2 were generated randomly using the code `cooperate.py` in the electronic supplementary material.

Authors' contributions. O.P. suggested the study. O.P. and A.A. conducted it and wrote the manuscript.

Conflict of interest declaration. The authors declare that they have no competing interests.

Funding. The authors are grateful to Baillie Gifford, The Mercers' Company, J. Anderson, M. Elsberg, W. Miller, F. Mustafa and C. Vermont for supporting the Ergodicity Economics programme at the London Mathematical Laboratory.

Acknowledgements. The authors thank M. Buchanan, D. Krakauer and A. Lightner for helpful comments.

References

1. Gibran K. 1926 *The Prophet*. New York, NY: Alfred A. Knopf.
2. Nowak MA. 2006 Five rules for the evolution of cooperation. *Science* **314**, 1560–1563. (doi:10.1126/science.1133755)
3. Aktipis A. 2016 Principles of cooperation across systems: from human sharing to multicellularity and cancer. *Evol. Appl.* **9**, 17–36. (doi:10.1111/eva.2016.9.issue-1)
4. Peters O. 2019 The ergodicity problem in economics. *Nat. Phys.* **15**, 1216–1221. (doi:10.1038/s41567-019-0732-0)
5. Roper M, Dayel MJ, Pepper RE, Koehl MAR. 2013 Cooperatively generated stresslet flows supply fresh fluid to multicellular choanoflagellate colonies. *Phys. Rev. Lett.* **110**, 228104. (doi:10.1103/PhysRevLett.110.228104)
6. Short MB, Solari CA, Ganguly S, Powers TR, Kessler JO, Goldstein RE. 2006 Flows driven by flagella of multicellular organisms enhance long-range molecular transport. *Proc. Natl Acad. Sci. USA* **103**, 8315–8319. (doi:10.1073/pnas.0600566103)
7. McCloskey DM. 1976 English open fields as behavior towards risk. *Res. Econ. Hist.* **1**, 124–170.
8. Gillespie JH. 1977 Natural selection for variances in offspring numbers: a new evolutionary principle. *Am. Nat.* **111**, 1010–1014. (doi:10.1086/283230)
9. Lande R. 2007 Expected relative fitness and the adaptive topography of fluctuating selection. *Evolution* **61**, 1835–1846. (doi:10.1111/evo.2007.61.issue-8)
10. Lewontin RC, Cohen D. 1969 On population growth in a randomly varying environment. *Proc. Natl Acad. Sci. USA* **62**, 1056–1060. (doi:10.1073/pnas.62.4.1056)
11. Orr HA. 2009 Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.* **10**, 531–539. (doi:10.1038/nrg2603)
12. Winterhalder B. 1986 Diet choice, risk, and food sharing in a stochastic environment. *J. Anthropol. Archaeol.* **5**, 369–392. (doi:10.1016/0278-4165(86)90017-6)
13. Capilla-Lasheras P, Harrison X, Wood EM, Wilson AJ, Young AJ. 2021 Altruistic bet-hedging and the evolution of cooperation in a Kalahari bird. *Sci. Adv.* **7**, eabe8980. (doi:10.1126/sciadv.abe8980)

14. Uitdehaag JC. 2011 Bet hedging based cooperation can limit kin selection and form a basis for mutualism. *J. Theor. Biol.* **280**, 76–87. (doi:10.1016/j.jtbi.2011.04.002)
15. Starrfelt J, Kokko H. 2012 Bet-hedging – a triple trade-off between means, variances and correlations. *Biol. Rev.* **87**, 742–755. (doi:10.1111/brv.2012.87.issue-3)
16. Schreiber SJ. 2015 Unifying within- and between-generation bet-hedging theories: an ode to J. H. Gillespie. *Am. Nat.* **186**, 792–796. (doi:10.1086/683657)
17. Kennedy P, Higginson AD, Radford AN, Sumner S. 2018 Altruism in a volatile world. *Nature* **555**, 359–362. (doi:10.1038/nature25965)
18. Hamilton WD. 1964 The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–16. (doi:10.1016/0022-5193(64)90038-4)
19. Aitchison J, Brown JAC. 1957 *The lognormal distribution*. Cambridge, UK: Cambridge University Press.
20. Redner S. 1990 Random multiplicative processes: an elementary tutorial. *Am. J. Phys.* **58**, 267–273. (doi:10.1119/1.16497)
21. Meadows DH, Meadows DL, Randers J, Behrens III WW. 1972 *The limits to growth*. New York, NY: Universe Books.
22. van Kampen NG. 1992 *Stochastic processes in physics and chemistry*, 3rd edn, Amsterdam, The Netherlands: Elsevier.
23. Leigh Jr EG. 2010 The group selection controversy. *J. Evol. Biol.* **23**, 6–19. (doi:10.1111/jeb.2009.23.issue-1)
24. West SA, Griffin AS, Gardner A. 2007 Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20**, 415–432. (doi:10.1111/jeb.2007.20.issue-2)
25. Darwin C. 1859 *On the origin of species by means of natural selection*. London, UK: John Murray.
26. Abbot P *et al.* 2011 Inclusive fitness theory and eusociality. *Nature* **471**, E1–E4. (doi:10.1038/nature09831)
27. Nowak MA, Tarnita CE, Wilson EO. 2010 The evolution of eusociality. *Nature* **466**, 1057–1062. (doi:10.1038/nature09205)
28. Wilson DS, Sober E. 1994 Reintroducing group selection to the human behavioral sciences. *Behav. Brain Sci.* **17**, 585–608. (doi:10.1017/S0140525X00036104)
29. Wynne-Edwards VC. 1986 *Evolution through group selection*. Oxford, UK: Blackwell Scientific.
30. Berman Y, Peters O, Adamou A. 2021 Wealth inequality and the ergodic hypothesis: evidence from the United States. *J. Income Distrib.* (doi:10.25071/1874-6322.40455)
31. Stojkoski V, Utkovski Z, Basnarkov L, Kocarev L. 2019 Cooperation dynamics in networked geometric Brownian motion. *Phys. Rev. E* **99**, 062312. (doi:10.1103/PhysRevE.99.062312)
32. Yaari G, Solomon S. 2010 Cooperation evolution in random multiplicative environments. *Eur. Phys. J. B* **73**, 625–632. (doi:10.1140/epjb/e2010-00027-4)
33. Peters O, Adamou A. 2018 The time interpretation of expected utility theory. (<http://arxiv.org/abs/1801.03680>)
34. Peters O, Adamou A. 2022 The ergodicity solution of the cooperation puzzle. Figshare. (<https://doi.org/10.6084/m9.figshare.c.5885340>)