

Research Article

Complementary Variety: When Can Cooperation in Uncertain Environments Outperform Competitive Selection?

Martin Hilbert

Department of Communication, University of California, Davis, Kerr Hall 369, Davis, CA 95616, USA

Correspondence should be addressed to Martin Hilbert; hilbert@ucdavis.edu

Received 9 February 2017; Revised 22 May 2017; Accepted 5 June 2017; Published 7 September 2017

Academic Editor: Sergio Gómez

Copyright © 2017 Martin Hilbert. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evolving biological and socioeconomic populations can sometimes increase their growth rate by cooperatively redistributing resources among their members. In unchanging environments, this simply comes down to reallocating resources to fitter types. In uncertain and fluctuating environments, cooperation cannot always outperform blind competitive selection. When can it? The conditions depend on the particular shape of the fitness landscape. The article derives a single measure that quantifies by how much an intervention in stochastic environments can possibly outperform the blind forces of natural selection. It is a multivariate and multilevel measure that essentially quantifies the amount of complementary variety between different population types and environmental states. The more complementary the fitness of types in different environmental states, the proportionally larger the potential benefit of strategic cooperation over competitive selection. With complementary variety, holding population shares constant will always outperform natural and market selection (including bet-hedging, portfolio management, and stochastic switching). The result can be used both to determine the acceptable cost of learning the details of a fitness landscape and to design multilevel classification systems of population types and environmental states that maximize population growth. Two empirical cases are explored, one from the evolving economy and the other one from migrating birds.

1. Introduction

According to Darwin, natural selection “is as immeasurably superior to man’s feeble efforts, as the works of Nature are to those of Art” [1, p. 65]. Nevertheless, we often have the ambition to go beyond the contributions competitive selection among different types has to offer. We know of a large variety of redistribution strategies in nature and society that aim at creating fitter populations. For example, in economic evolution, conscious portfolio managers can increase the growth rates of the entire market by constantly reshuffling resources despite current market selection [2–5]. In the political economy, countercyclical fiscal policies have been a common policy tool over the last century [6, 7]. And, in biological populations, stochastic phenotype switches redistribute phenotypes among offspring to offset the acting forces of selection [8–12].

The details of the implementation of those redistribution mechanisms might be diverse but do not matter for our purposes. We ask under which conditions it is possible that an interventionist resource distribution strategy among types

can outperform blind competitive selection among types in a stochastic environment. What are the conditions under which, for the sake of population’s growth, it is more efficient to relocate resources among its individuals than to let nature and markets do the selection? We answer the question by coining a new term: complementary variety. We find that it is only beneficial for members of a population to cooperate if different population types complement each other’s fitness in different environmental states. We also see that, with complementary variety, it is always beneficial to cooperate in the light of an uncertain future environment. Complementary variety provides the *sine qua non* boundary condition for beneficial resource redistributing cooperation in stochastic environments. The contribution of the article consists in providing a single measure that quantifies the complementary variety of a population in a specific environment and linking it to the potential gain of strategic intervention.

1.1. The Fitness Gains of Intervention. For the case of a constant environment (a single environmental state), during

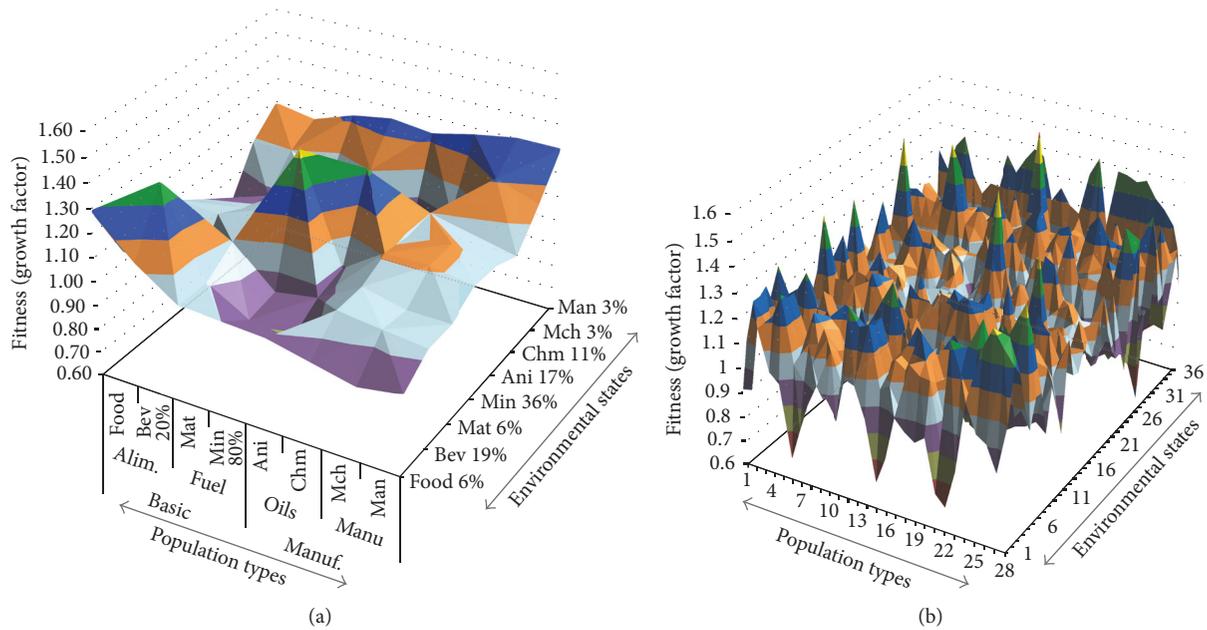


FIGURE 1: Fitness landscape where the bottom dimension refers to different population types (groups of the overall population) and environmental states (temporal periods), while the height represents fitness (the growth rate of each type in each state, following the traditional definition of fitness in fitness landscapes (28)). Both (a) and (b) represent the same case, while (a) is coarse-grained into 8 population types and 8 environmental states; and (b) is fine-grained into 28 population types and 36 environmental states (based on data from [22–24]).

which population types have stable relative growth rates, the best resource distribution strategy always simply comes down to placing all resources on the fastest growing type. If the fitness matrix does not change as a result of this intervention (no density dependence), the population fitness after this strategic intervention will be equal to the growth rate of the fittest type, which is always at least as large as the population fitness eventually attainable through hands-off natural selection. In the first case, we say that types use strategic cooperation (W_{SC}) to maximize overall population growth, while, in the second case, we talk about hands-off competitive selection (W_{CS}). In order to obtain a gain from strategic cooperation, we need to identify the growth rate of each type/group and make the effort to redistribute resources to the type with maximal fitness. This might have a cost, but as long as the obtained fitness gain of the intervention ($W_{SC} - W_{CS}$) is larger than the involved cost, it is still worthwhile. We are interested in the potential benefit of this sort.

What about fluctuating environments during which the growth rates of different members change over time? For example, Figure 1(a) shows a fitness landscape in which types have different fitness values in different environmental states that appear with different probability. Our first research question relates to such cases:

RQ1: When can strategic cooperation outperform competitive selection in stochastic environments?

In terms of the previously introduced notation, we ask for a measure of the fitness matrix that allows us to quantify when a resource redistribution strategy in fluctuating environments is worthwhile: $W_{SC} - W_{CS} > 0$. This quantity also gives us an

upper bound for the maximal cost for which it is profitable to obtain information about the fitness landscape and on executing a resource distribution mechanism. Only if there is a potential benefit for intervention do we need to study the landscape. Otherwise we just let blind selection run its course.

As a general lead, we know the optimal strategy or, in the words of Richard Levins, the “correspondence between the optimal structures of populations and species and the pattern of environmental heterogeneity in space and time” [13]. It is provided by the literature of bet-hedging [2, 9–12, 14–18], also known as stochastic switching in biology [8, 19] and log-optimal portfolios in economics [3–5, 20] (see Supporting Information S4, in Supplementary Material available online at <https://doi.org/10.1155/2017/5052071>, for an introduction to this literature). This literature shows that in some cases one can increase population fitness beyond the growth rate obtainable by any single type by spreading resources among different types. This is done by constantly redistributing resources back to the same (optimized) distribution of resources among types, counteracting the up and down of the changes produced by the varying fitness landscape. It can be shown that bet-hedging achieves optimal long term population growth [21]. In the worst case (under some mathematical conditions referred to as being “outside the region of bet-hedging”), bet-hedging optimization suggests betting all resources to the fastest growing type, which is equivalent to the optimal strategy for unchanging environments [10, 14–16].

Unfortunately, the literature of bet-hedging is surprisingly quiet about the general boundary conditions under which it makes sense to apply it. RQ1 explores these boundary

conditions and we generalize it to even nonoptimized strategies. With “strategic cooperation” we refer to any redistribution of resources that holds population shares constant over time. In a stochastic environment with uncertainty about the realization of the next environmental state, that is the best one can do. In an uncertain stochastic environment, we know the distribution, and during our strategy we might sometimes win and sometimes lose as lady luck rolls the environmental dice. However, we aim at winning on average, holding on to our strategy in every turn. In the final discussion we relate our generalized results back to the literature of optimized bet-hedging, which turns out to be a special case of our result. Our results will therefore provide the conditions *sine qua non* for the applicability of bet-hedging.

1.2. The Fitness Gains of Coarse-Graining. The foregoing question leads to the related question when it is worthwhile for the population to undergo the effort of working with a more fine-grained fitness landscape in space and time. If it pays off to allocate resources on the fitter type in a constant environment, then it is also always beneficial to fine-grain the population by splitting types into subgroups. This is because—per definition—the growth rate of a more aggregate higher level type is the average between subtypes of which at least one has to have higher fitness than the average (or at least all equal). We can then allocate all resources on this even fitter lower level type. This shows that the question of optimal intervention is inherently linked to the level of coarse-graining of population types.

For example, both Figures 1(a) and 1(b) show the same empirical evidence over the same period, with the difference that the latter distinguishes among almost 16 times more points in the fitness landscape, while the former averages over these details. The 36 environmental states and 26 population types of Figure 1(b) are coarse-grained into 8 states and 8 types in Figure 1(a).

The question is equivalent to asking when it pays off to learn the detailed particularities of a fitness landscape to execute resource redistribution strategy among more fine-grained subtypes (such as Figure 1(b)) and when it is enough for the population to simply consider a more coarse-grained outlook (such as Figure 1(a)), which implies letting the “blind forces” of natural selection “ride the waves” of the lower levels of the landscape.

This question is essentially the same as our previous question and comes down to quantifying the potential benefit obtained by a strategic intervention on the more detailed fitness landscape. If $W_{SC} - W_{CS} < 0$, we might as well let competitive selection ride over the more detailed peaks and valleys of the fitness matrix in a blind manner.

In practice we see that fine-graining seems to pay off, as we observe that both biological and social evolution seem to distinguish among ever more detailed characteristics. Throughout history, the increasing number of branches of the tree of life has grown life’s diversity [26], and the evolving economy, once made of a handful of generic professions, adds new job types much quicker as it takes old ones away, evolving an ever more fine-tuned division of labor [27]. At the same time, evolving populations also seem to distinguish

among more fine-grained environmental states. Animals that learn how to read weather situations in more detail increase survival chances [28], and data analytic companies that distinguish among ever more detailed market conditions maximize sales [29, 30]. While both processes seem to be general tendencies, it is not clear which one of them is more beneficial and how they relate to each other. We can analyze these questions by applying the newly developed methodological toolkit from RQ1:

RQ2: How much fitness can potentially be gained by disaggregating the fitness landscape in more detailed population types and/or in more detailed environmental states?

1.3. Outline. The following section answers RQ1 by deriving our theoretical toolkit. As a first step, we compare the achievable fitness of competitive selection and strategic cooperation with help of the more common analytical frameworks of the variance and covariance of fitness. This is useful to build basic intuition but is limited to simple cases. In order to advance to more realistic multivariate cases, we subsequently analyze the same setup with the help of a measure from information theory [21, 31]. We measure selection forces with the Kullback-Leibler divergence (so-called relative entropy). This also allows us to derive a straightforward way to analyze multilevel dynamics over nested levels of coarse-graining in a single equation.

We then continue to explore RQ2 by applying the developed measures to two empirical examples from biology and the evolving economy. We analyze in which cases it would increase the fitness of the population of migrating birds in North America if they would select their temporal homes according to more detailed geographical locations and if the export portfolio of Japan’s economy would benefit from creating a trade strategy that goes beyond a coarse-grained distinction of manufactured and nonmanufactured goods and would make more detailed distinctions among subtypes such as machinery, chemicals, and oils.

In the Discussion, we link the identified concept to the well-established literature on bet-hedging and review both limitations and arising research questions in the conclusions.

2. Theoretical Methods: RQ1

The evolving population is subdivided into different population groups G , with each group-type g consisting of a certain number of discrete units that appear with frequency $p(g)$. Types can be identified by any exclusive and exhaustive variable or property of a group, such as allele, genotype, phenotype, location, attribute, characteristic, or community membership. The application of the resulting equations does not depend on the choice of any ecological, morphological, physiological, molecular, behavioral, geographic, sociodemographic, or any other way to partition the population into types.

Population fitness $\overset{g}{W}$ is the factor of reproduction or growth of the total number of population units N . The population fitness indicates the average number of offspring:

TABLE 1: Fitness landscapes in matrix form with types g and environmental states e and corresponding fitness values $w(g, e)$. The larger fitness value of the binary type split (horizontal) is denoted with capital letter W and the less fit of the two with the small letter w . Despite this simple case of two types, the covariance relation of (2) also holds also for more than two types. (a) schematizes a positive covariance between type fitness in environments $e = \alpha$ and $e = \beta$, while (b) schematizes a negative covariance, with superior type fitness along the diagonal of the matrix.

(a)		
$w(g, e)$	Population type level	
Environ. state	$g = A$	$g = B$
$e = \alpha$	$W(A, \alpha)$	$w(B, \alpha)$
$e = \beta$	$W(A, \beta)$	$w(B, \beta)$

(b)		
$w(g, e)$	Population type level	
Environ. state	$g = A$	$g = B$
$e = \alpha$	$W(A, \alpha)$	$w(B, \alpha)$
$e = \beta$	$w(A, \beta)$	$W(B, \beta)$

$\overset{g}{W} = N^{+1}/N$, wherein the superscript +1 indicates the updated generation after reproduction (a shorthand for the long-hand notation $t + 1$). The top-script g indicates that this is the average growth rate over all types G . The fitness per population type is $w(g)$. Total population fitness is the expected value of the fitness of each group type: $\overset{g}{W} = \sum_g p(g) * w(g) = E[w(g)]$.

If the environment changes, types can have different fitness values in different environmental states. In a stationary time series, each environmental state has a certain probability of occurrence and can be represented by a random variable with distribution $P(E)$ (for simplicity assumed to be discrete). The fitness of type g in each environmental state e is represented with the following: $w(g, e)$ (see Table 1 for the basic setup of a fitness matrix between types and environmental states).

Much in the sense of game theory, we ask about potential fitness benefits of strategic cooperation if resources can be redistributed among types. Resource redistribution converts units from one type into units of another type. Such interventionist strategies are not possible for all evolving populations (genetic material cannot easily be “redistributed”). It is straightforward to apply such strategy to social populations in which the population units N represent redistributable resources (like money). It is also applicable to biological populations where a genotype can reproduce a nonproportional share of phenotypes (so-called “stochastic switching” [8–12]) and when organisms with cognitive abilities change their behavior according to different environmental patterns.

The constant redistribution of strategic intervention implies counteracting the forces of competitive natural selection to persistently return to the same stable portfolio distribution. In a varying environment, this means that the population before updating in any environmental state is the same ($P(G | e) = \text{constant for all } e$). For our comparative

purposes between the mean population fitness of strategic cooperation (W_{SC}) and competitive selection (W_{CS}), we set the initial population distribution for the case of natural selection equal to the stable portfolio population distribution for the case of strategic intervention: $P^0(g) = P(G | e)$ for all e . This makes it straightforward to compare both cases directly with each other.

2.1. One- and Two-State Environments: Variance and Covariance. We start by analyzing the two simple cases of unchanging environments and of stochastic environments with two states. We can analyze both in terms of more common variance and covariance measures.

For unchanging environments where the fitness of each type is unchanging, Fisher’s fundamental theorem of natural selection applies. It equates the rate of increase in mean population fitness to the variance in fitness of its types [32]. After about half a century of much confusion in the literature, it was finally clarified that Fisher meant his theorem only to apply to this case of unchanging type fitness, as only in this case can natural selection select consistently among types [33–35]. What is sometimes confusing is that Fisher regarded all possible exogenous and endogenous reasons for changing type fitness as a matter of the environment (such as dominance, density dependence, epistasis, climate, and interactions with other species). He therefore defined the state of fixed type fitness as one single (unchanging) environmental state: “what Fisher’s theorem tells us is that natural selection [...] at all times acts to increase the fitness of a species to live under the conditions that existed an instant earlier” [35, p. 131].

In this case, any strategic cooperation that holds shares of types constant over time can never lead to higher population fitness than competitive selection. By artificially holding the shares of types constant, selection cannot contribute to the increase of population fitness by increasing the share of the fitter types:

$$\overset{g}{W}_{CS} \geq \overset{g}{W}_{SC} \longleftrightarrow \quad (1)$$

$$\text{var}(w(g)) \geq 0.$$

The best one can do is to equate both by mimicking the end-result of competitive selection (only the fittest type survived) and place all resources on the fittest type (equation [S5] in Supporting Information).

If there are two different environmental states, the role of the variance can sometimes be represented as a covariance of type fitness in both environmental states. This applies to a special case where both environmental states occur with equal probability (in line with Table 1, where $p(e = \alpha) = p(e = \beta)$).

It turns out that no strategic cooperation can be beneficial if the fitness values among different types in different environmental states do not complement each other. Complementarity means that some types grow faster in some environmental states, while other types grow faster in other environmental states (on average). We can use the covariance of type fitness in both environmental states to quantify the

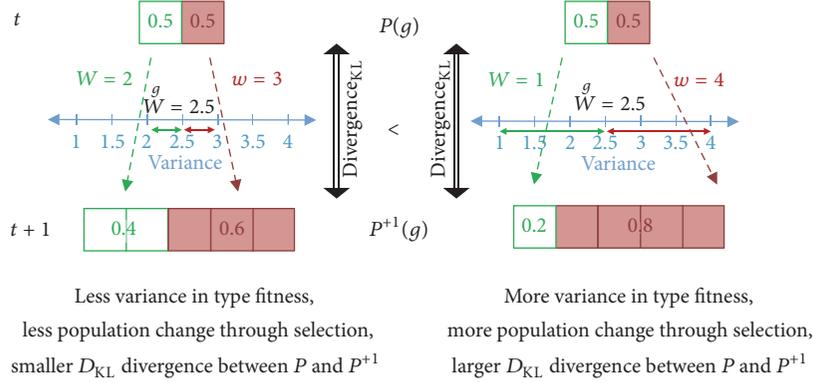


FIGURE 2: Illustration of Kullback-Leibler relative entropy as a measure of strength of evolutionary selection. The left hand side illustrates a case with less variance in type fitness, less population change through selection, and smaller D_{KL} divergence between P (0.5; 0.5) and P^{+1} (0.4; 0.6). The right-hand side illustrates the case with larger variance in type fitness, more population change through selection, and larger D_{KL} divergence between P (0.5; 0.5) and P^{+1} (0.2; 0.8).

average tendency of type fitness in both environmental states. Table 1(a) provides an illustrative sketch of a positive correlation: here type A is the fitter type in both environmental states α and β . In this case there is no complementary variety and it is not possible to increase population fitness by a strategic cooperation. This also means that it is not even worthwhile to distinguish between those two environmental states $e = \alpha$ and $e = \beta$. We might as well simply let natural selection run its course and not worry about the environmental distribution and its more fine-grained fitness landscape. The average population fitness achieved by competitive selection ($\bar{W}_{CS}^{e,g}$) will always be greater than or equal to the population fitness achieved by a strategic cooperation ($\bar{W}_{SC}^{e,g}$) (Supporting Information S3.1):

$$\bar{W}_{CS}^{e,g} \geq \bar{W}_{SC}^{e,g} \iff \text{cov}(w(g, e = \alpha), w(g, e = \beta)) \geq 0, \quad (2)$$

where \bar{W} is the geometric mean of the population fitness over environmental states: $\bar{W} = \prod_e (\bar{W}^{p(e)})^{e,g}$.

This shows that a negative covariance is the necessary condition for a potential benefit of strategically cooperating among types by holding the type shares constant (in the optimal case through portfolio theoretic bet-hedging). The schematization of a negative covariance in Table 1(b) visualizes that this implies that there is a complementary variety among the different types in the two environmental states: one type that is more specialized for one environmental state and the other type for the other state.

The covariance measure provides useful intuition for the concept of complementary variety of a fitness landscape. The problem with it is that it can only handle two covarying environmental states. We will therefore develop the same ideas by using the information theoretic measure of relative entropy, which allows quantifying the same logic for any kind of multivariate environmental distribution with any number of stochastically varying environmental states.

2.2. Multistate Environments: Relative Entropies. When working with information theoretic decomposition of fitness we usually represent fitness on a logarithmic scale. This does not change the generality of the results. Logarithmic growth factors are the norm in economics and in biology often referred to as Malthusian fitness [32, 36, 37]. It is useful since fitness “adds up” through multiplication over time ($\prod w$) and the logarithm converts this geometric process into a more intuitive linear summation ($\log(\prod w) = \sum \log(w)$). We will use logarithms of base 2, which represents fitness in terms of the number of population doublings at each time step and the arising informational quantities in bits.

2.3. Single Step Updating. We start by expressing a simple replicator step in terms of relative entropy. The proportion of a type after reproduction is defined by the so-called “replicator equation”: $p^{+1}(g) = p(g) * w(g) / \bar{W}$. The log population fitness can be decomposed in the following way (Supporting Information S2):

$$\log(\bar{W}) = E[\log w(g)] + D_{KL}(P(g) \parallel P^{+1}(g)), \quad (3)$$

where $E[\dots]$ stands for the expected value of population types ($E[\dots] = \sum_g p(g) * [\dots]$) and D_{KL} is the relative entropy or the Kullback-Leibler divergence: $D_{KL}(P(g) \parallel P^{+1}(g)) = \sum_g p(g) * \log(p(g)/p^{+1}(g))$ [21, 38]. The relative entropy D_{KL} is perhaps the most general measure of information theory and has Shannon’s famous absolute entropy and mutual information as special cases (for a short primer on D_{KL} , see Supporting Information S1). It has some insightful information theoretic interpretations as the inefficiency of an average code (see S1) but is essentially a directed measure of the degree of divergence between two distributions, in our case P and P^{+1} (see Figure 2). D_{KL} is always nonnegative and only zero if there is no change in the population, with $P = P^{+1}$ [21].

As such, D_{KL} can be used as a measure that quantifies the strength of evolutionary selection [39–42]. As shown by Frank [43–45], D_{KL} can be related to the variance in

type fitness and therefore to Fisher's fundamental theorem of natural selection [32]. The larger the variance among the fitness of different types, $\text{var}[w(g)]$, the larger the difference between $P(G)$ and $P^{+1}(G)$ and the larger $D_{\text{KL}}(P(g) \parallel P^{+1}(g))$. Figure 2 visualizes this relation.

With this in mind, (3) says that population fitness is large when the expected value term is large and the D_{KL} term large. The expected value is taken over the updated generation P^{+1} and is therefore large when the fitter types increase and the less fit ones decrease their shares, while D_{KL} is small when the population has reached equilibrium and does not change anymore (being zero with no divergence, at $P^{+1} = P$). Equal to our previous argument for unchanging environments, this provides another way of showing that the mean population fitness $\overset{g}{W}$ is at its maximum if we move all resources to the fittest type. In this case the expected value term is at its maximum and the D_{KL} term at its minimum.

2.4. Average Updating. When considering multiple stochastic environmental states, we apply the decomposition of (3) to the joint probability distributions between the population and the environment, $P(G, E)$ [46]. This leads to the following version of our previous equation (see Supporting Information S2):

$$\begin{aligned} \log \overset{e,g}{W} &= E_{e,g} [\log w(g, e)] \\ &+ D_{\text{KL}}(P^-(g | e) \parallel P^+(g | e)). \end{aligned} \quad (4)$$

A subtle but very important detail of this decomposition is that in this case P^- and P^+ refer not to the initial and final distribution of the entire observational period (the superscript + is not equal to the superscript +1 we used before) but to the average distribution before and after typical updating over these different environmental states.

For all practical purposes, these average distributions can be calculated in the following way. $P(E)$, $\overset{g}{W}(e)$, and $w(g, e)$ can be detected empirically by measuring the fitness values as geometric means during a given environmental state. This defines the average population distributions in a certain environment, $P^-(G | e)$, since $\overset{g}{W}(e) = \sum_g p(g | e) * w(g, e)$. Based on this result we can calculate the average population share after average updating in a certain environment by using the replicator equation: $p^+(g | e) = p^-(g | e)w(g, e)/\overset{g}{W}(e)$. Note that both resulting distributions $P^-(G | E)$ and $P^+(G | E)$ represent average distributions which depend both on the fitness matrix and also on the timeframe of evolutionary observation (which affects the average population fitness per environmental state $\overset{g}{W}(e)$). Finally, we can calculate the joint distributions $P^-(G, E)$ and $P^+(G, E)$ by multiplying the conditional distributions per environmental state with the observed environmental distribution $P(E)$. The relative entropy D_{KL} in (4) works with these joint and conditional distributions.

The question of whether competitive selection is potentially superior to strategic cooperation turns out to hinge on

the superiority of the arithmetic mean of the geometric mean or, vice versa, the geometric mean of the arithmetic mean:

$$\overset{e,g}{W}_{\text{CS}} = E_g \left[\prod_e w(g, e)^{t_e} \right] \geq \text{or} \leq \prod_e (E_g [w(g, e)])^{t_e} = \overset{e,g}{W}_{\text{SC}}, \quad (5)$$

where $E_g[\dots]$ stands for the expected value over all types and t_e for the number of periods of occurrence of environmental state e during the timeframe of observation. It turns out that this question is equivalent to the question of whether the average relative entropy between the first and the last periods of evolutionary observation is proportionally larger or smaller than the joint relative entropy during average updating. We quantify this with

$$\Delta D_{\text{KL}} = T * D_{\text{SC}}(P^- \parallel P^+) - D_{\text{CS}}(P^0 \parallel P^T). \quad (6)$$

The term D_{CS} is the Kullback-Leibler divergence that measures the average strength of selection between the first and the last period of the entire observational period, while in between natural selection runs its course. P^0 stands for the distribution of the initial population and P^T for the final distribution of the population after the total number T of observed periods of the evolutionary dynamic ($T = \sum_e t_e$). This is much in line with our (3), since we simply look at population change between two points in time. The term D_{SC} is in line with (4). It quantifies the average updating per period that is counteracted by a resource distribution strategy. Since D_{SI} measures the average updating per period and D_{CS} the span over all periods, the former is multiplied with the total number of periods T .

Supporting Information S3.2 shows ΔD_{KL} is proportional to the potential increase in population fitness $\overset{e,g}{\Delta W} = \overset{e,g}{W}_{\text{SC}} - \overset{e,g}{W}_{\text{CS}}$ or $\overset{e,g}{\Delta W} \propto \Delta D_{\text{KL}}$. Written out in full, we get

$$\begin{aligned} \left[\overset{e,g}{W}_{\text{SC}} - \overset{e,g}{W}_{\text{CS}} \right] &\propto \left[T * D_{\text{SC}}(P^-(g | e) \parallel P^+(g | e)) \right. \\ &\left. - D_{\text{CS}}(P^0(g) \parallel P^T(g)) \right]. \end{aligned} \quad (7)$$

This quantifies the level of complementary variety in the fitness landscape. To see this, note that, during each new period (and therefore at each environmental state e), strategic cooperation assures that $p^-(g | e)$ is held constant at the same distribution as was set at the beginning ($P^-(G | e) = P^0(G)$ for all e). Based on this stable portfolio distribution, $P^+(g | e)$ is calculated on the basis of the replicator equation and therefore considers the occurring selection forces despite any intervening redistribution. This is justified because, even with a proactive strategy that keeps shares constant, some types effectively increase and others decrease their share temporally in a particular environmental state. For example, a portfolio manager on the stock market has real gains and losses as market selection effectively acts on the portfolio but constantly counteracts them by constant intervention that counterbalances market selection in order to keep the portfolio stable. D_{SC} quantifies the average strength of selection that is counteracted by the strategy.

With complementary variety in the fitness landscape, the fitter types in one environment are mainly the less fit

types in the other environment, and selection works in opposing directions in different environments. As a result, growth tendencies in the different environmental states rather balance each other out by shifting gains back and forth. There will be relatively little change between the population distributions of the first and the last periods. The divergence $D_{CS}(P^0 \parallel P^T)$ will be relatively small. However, the fact that selection conditioned on one environmental state will counteract selection in another environmental state increases the aggregated conditional divergence over all environmental states: $D_{SC}(P^- \parallel P^+)$ (for more on conditional Kullback-Leibler divergence see Supporting Information S1). The counterbalancing tendencies represent complementary variety.

In this sense, in line with the interpretation of D_{KL} as a measure of the strength of selection (see Figure 2), it can be used as a nonnegative multivariate measure to quantify the amount of complementary variety among fitness values in different environmental states. The more the fitness landscape is skewed to opposing directions in different environmental states, the larger the average D_{SC} divergence of different environmental states is and the proportionally larger the potential benefit of strategic cooperation is.

The surprising finding of (6) is that the existence of complementary variety assures that any population distribution that is held constant will always increase population fitness in comparison to competitive selection. It is not necessary to calculate the optimal population distribution in the sense of an optimized bet-hedging strategy. The law of complementary variety states that any intervention that holds population shares constant will always outperform hands-off natural selection, if the environmentally conditioned relative entropy over all periods is larger than the relative entropy between the first and last period.

2.5. Coarse-Graining of Types and States. We now link our new results to the previously introduced question of the benefit of coarse-graining. One of the benefits of working with relative entropy D_{KL} is that it naturally lends itself to multilevel decomposition. This is done through the chain rule of relative entropies [21], which exploits the nestedness of conditional probabilities.

The question of course-graining applies to both population types and environmental states. For both we track the levels of coarse-graining with a subscript index $L = \{0, 1, 2, 3, \dots, i\}$, in which the first level $l = 0$ refers to the most coarse-grained aggregation.

For the coarse-graining of types, G_0 represents the entire population and the lowest level $l = i$ refers to the most fine-grained level of indivisible individuals, with the set G_i containing N different members. The intermediate level consists of nested subdivisions of higher level types into lower level types, $g_{l+1} \in g_l$. This does not require that each higher level type consists of the same number of lower level types. In other words, levels consist of nested hierarchies, where a lower level is some kind of subdivision of higher level types into lower level types. The shares of the types in lower levels, g_2 , are conditioned on a higher level type, g_1 : $p(g_2 \mid g_1)$. For the corresponding type fitness, nested conditioning does not have an effect and $w(g_2 \mid g_1) = w(g_2)$. In this

sense, unconditioned higher level fitness is the average of conditioned lower level constellations: $W(g_l) = \sum_{g_{l+1}} p(g_{l+1} \mid g_l) * w(g_{l+1})$.

For the coarse-graining of environmental states, the ultimate lowest level l depends on the smallest meaningful discretization of time. The highest level E_0 represents the entire period of observation. If we take the geometric mean of the growth rates of smaller periods, we can calculate one single average growth rate over all periods. Any subsequent lower level environmental states are the temporal subperiods that create any aggregated higher level geometric mean fitness: $w(g_1) = \prod_e w(g_2, e)^{p(e)}$ (see Supporting Information S3.3 for an example).

Using these two ways of coarse-graining types and states allows us to expand (4) over multiple levels. The logic is similar to Price's well-known recursive multilevel decomposition of evolutionary dynamics [43, 44, 47, 48]. We decompose type fitness in a recursive manner by applying the same logic with which we decomposed population fitness. Equation (8) works with joint probabilities of G and E . However, for readability, we use shorthand subscripts, such as $P_{2|1}$, to indicate the nested conditioning of the joint distribution $P(g_2, e_2 \mid g_1, e_1)$. For reasons of readability, we will focus on the population level, $l = 0$, different types, $l = 1$, and subtypes, $l = 2$, but in general this logic can be followed for any g_l and g_{l+1} . Using the same shorthand, (8) is the same as (4):

$$\log W(g_0) = \sum_{g_1, e_1} p_1^- * \{\log w_1\} + D_{KL}(P_1^- \parallel P_1^+), \quad (8)$$

$$\begin{aligned} \log W(g_0) &= \sum_{g_1, e_1} p_1^- \left\{ \sum_{g_{2|1}, e_{2|1}} p_{2|1}^- * \{\log w_{2|1}\} + D_{KL}(P_{2|1}^- \parallel P_{2|1}^+) \right\} \\ &\quad + D_{KL}(P_1^- \parallel P_1^+). \end{aligned} \quad (9)$$

The term in the curly brackets decomposes each type of fitness from level 1 into conditioned lower level types on level 2 by applying a conditional version of the same equation. This results in a D_{KL} term for the higher level and (a conditional) one for the lower level.

We can now apply the chain rule for relative entropy to (9), which brings us back to the form of (4), just on the more fine-grained level (see Supporting Information S1):

$$\begin{aligned} \log W(g_0) &= \sum_{g_{1,2}, e_{1,2}} p_{1,2}^- \{\log w_{1,2}\} + D_{KL}(P_{1,2}^- \parallel P_{1,2}^+) \\ &= \sum_{g_2, e_2} p_2^- \{\log w_2\} + D_{KL}(P_2^- \parallel P_2^+). \end{aligned} \quad (10)$$

The nested nature of hierarchies results in the fact that the sum of the conditional relative entropies over all levels (see (9)) is equal to the unconditioned relative entropy on the lowest (most fine-grained) level of types (see (10)). We can always apply the chain rule from the highest level to the lowest discernable level of indivisible individuals, with maximally $\log_2(N)$ nested levels of types (since the smallest possible group size at each level is 2).

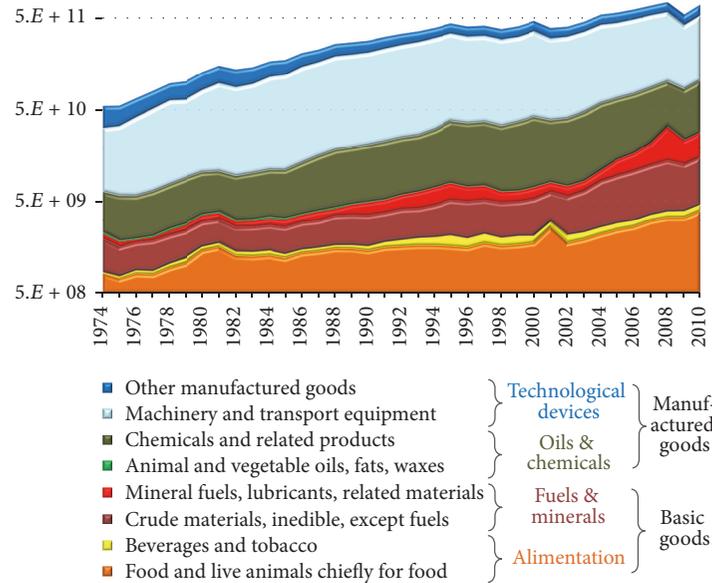


FIGURE 3: Japan's export products according to the first-digit classification of Standard International Trade Classification (SITIC), revision 2 for 1974–2010 in nominal value US\$ (based on [22–24]). The 8 main export groups of the first-digit classification are aggregated into two binary groupings on higher levels. The highest level distinguishes between basic goods and manufactured goods. Basic goods consist of alimentation and fuels and minerals and manufactured goods of oils and chemicals and manufactured devices (semilog plot).

Equations (9) and (10) allow us to explore multilevel trade-offs in strategic cooperation. Equation (9) proposes that we can intervene and redistribute at different levels. We can redistribute resources among types of level 1 and/or for subtypes within those types. This leads to a well-known problem to the literatures of both evolutionary biology and social science. On the one hand, cooperating on the lower level aims at increasing the fitness of a specific population, which ultimately aims at increasing the competitiveness of this population with other populations on the next higher level. We do strategic cooperating on the lower level in order to influence competitive selection on the higher level. On the other hand, strategic cooperation on certain levels sets constraints on lower levels. For example, if governments cross-subsidize among industries or sectors, it does constrain market forces on the firm level and potentially hamper the emergence of disruptive innovations. In practice, the viability of any strategic cooperation among lower level types (see (10)) depends on the reach of resource redistribution flexibility. In some cases resource redistribution might be restricted to a particular higher level type (within an economic sector or genetic family).

3. Empirical Results: RQ2

Equipped with the theoretical toolbox that we developed to answer RQ1, we now explore RQ2 with two empirical cases. The goal is to explore the involved orders of magnitude of fitness gains arising from strategic cooperation. We do so by calculating the potential gains of optimized strategic cooperation for a successively fine-grained fitness matrix given by empirical data. We then quantify fitness gained by distinguishing more detailed population

types and by considering more fine-grained environmental states.

The first case is taken from the evolving economy, that is, Japan's trade export (data taken from [24], that is, the cleaned version of [22, 23]) (see Figure 3). The second case uses the data provided by the North American Breeding Bird Survey and divides the population up into geographic locations of birds [25]. It is important to emphasize that the idea of the following applications is not to learn anything new about Japanese exports or North American birds but to study the behavior of our equations when applied to real-world data.

This leads us to the question of how to define population types and environmental states. The United Nations Statistical Commission reports working on the basis of 870 different national classifications from 154 countries to identify economically relevant types [49]. The question of meaningful environmental states is even trickier. An environmental state can be defined as any recurring characteristic in time, such as the weather, the state of the global economy, or seasons, such as “every other year.”

This being said, we can use our decomposition for two goals. We can use it descriptively, by working with some arbitrarily chosen criteria for types and states, and then calculate how much we could improve fitness if we would strategically intervene. We can also use it normatively, by defining states and types in a way that aims at creating complementary variety. This means that we create environmental states in which different types complement each other and/or vice versa.

In the following, we do not explore the question of the best way of doing this, but we opt for a simple mixed approach. For the classification of types, we use traditional groupings from existing taxonomies: UN export classifications (see Figure 3) and geographic boundaries (see Figure 4).



FIGURE 4: Geographic coarse-graining of bird populations among 59 states in Northern America [25].

Based on these types, we then use a simplistic way to create a certain degree of complementary variety through the definition of corresponding environmental states: we define an environmental state as the state during which a particular type grows faster than any other type in the population. This can provide as many environmental states as types (one environmental state that is optimal for one particular type) but can also result in less environmental states than types (since a type could never be the fittest).

3.1. Japanese Exports. In 1974, the total population G_0 of the Japanese export economy counted US\$ 54.7 billion and in 2010 some \$701.1 billion. This results in a compound annual growth factor $\overset{e.g.}{W}$ of 1.07341 over 36 periods. We decompose this total population over three nested levels to the 8 main export groups of the official first-digit classification of UN COMTRADE SITC rev. 2 (see Figure 3).

The first row in Figure 5 successively fine-grains the total export population G_0 into finer types (from 2 types at the first level types G_1 , to 4 types at the second level G_2 , to eight types at G_3). Naturally, at each lower level we discover faster growing subtypes. Fitness optimization suggests placing all resources on the fastest growing type, which increases population fitness (see shaded types in Figure 5). We calculate the potential gain of this strategic intervention $\Delta\overset{e.g.}{W}$ as the difference between the reallocated resources and the empirically detected growth rate (competitive natural selection). For example, for the case “Space 2_Time 1” in Figure 5, $\Delta\overset{e.g.}{W} = \overset{e.g.}{W}_{IS} - \overset{e.g.}{W}_{NS} = 1.0738 - 1.07341 = 0.039\%$. This case is visualized in Figure 6(a).

We now consider respective environmental states. For a binary split of the Japanese export economy, we end up with two environmental states: one in which basic products grow faster than manufactured products (which happens in 15 of the 36 years in our sample or $p(e) = 0.42$) and vice versa for the other type (see Figure 5, “Space 2_Time 2”).

Fitness optimization over this 2-by-2 fitness matrix suggests a bet-hedging strategy that allocates a constant share of 55% of available resources to basic export products and 45% to manufactured products. This is visualized in Figure 6(b). The rest of the first column of cases in Figure 5 fine-grains the entire period into 4 and then 8 different environmental states. Figure 5 also shows the values for (4); that is, $\overset{e.g.}{W} = 2^{E[\log w]} * 2^{D_{KL}(P^* \| P^*)}$ (exponentiation gets rid of the logarithm).

Optimizing the fitness landscape with four population types and two environmental states (case “Space 4_Time 2” in Figure 5) suggests neglecting exports in alimentation and in technology products and holding a constant share of 33.8% of the resources for fuels and minerals and 66.2% for oils and chemicals. Going one level deeper and distinguishing among eight types suggests placing all resources on the subtype mineral fuels (“Space 8_Time 2” in Figure 5).

The rest of Figure 5 continues more fine-grained distinctions among population types and among lower level environmental states and mainly speaks for itself. Figures 6(c) and 6(d) visualize the cases with a square fitness matrix (equal number of types and states), with four types and states (Figure 6(c)) and eight types and states (Figure 6(d)). The values in Figure 6(d) are the same as in Figure 1(a), just discretized.

3.2. Migrating Birds. The following example asks about potential fitness gains if the North American population of migrating birds would distinguish among more fine-grained geographic locations. If the bird population sensed differences in growth potentials across North America and if it were free to reallocate to exploit stationary environmental patterns, how much could the population of migrating birds gain from more detailed distinctions?

The overall population (G_0) increased from 1,520,305 birds in 1980 to 1,896,379 birds in 2014. This results in a compound annual growth factor $\overset{e.g.}{W}$ of 1.00652 (geometric mean over 34 periods). The database provides migrating bird populations according to political states in North America [25]. Figure 4 shows the selected coarse-graining between north and south (referring to our first level types G_1), then east and west (adding the second level types G_2), and then a more fine-grained distinction between midnorth and midsouthern states to distinguish among 8 different third-level groups G_3 (roughly following different climate regions).

We follow the same logic as we did for the case of Japanese exports in Figure 5. On the first level we distinguish only two types in space (North and South of North America) and treat entire period of 34 years as one single updating reproductive cycle. Optimal growth with a binary distinction between north and south occurs when some 47% of the birds migrate to the Northern part (which is favorable 44% of the time). This strategic intervention contributes almost 1% of additional fitness to the population ($\Delta\overset{e.g.}{W} = 0.00969$ or about 15,000 additional birds). On the lowest level we distinguish among eight types in space and eight environmental states in time. Growth optimization here recommends that some 20% of the birds migrate to the Midnorthwest, while the remaining 80% should migrate to the Midnortheast (see Figure 4). This

→ Fine-graining of population types “in space” →

← Fine-graining of environmental states “in time” ←	Space 2_Time 1		Space 4_Time 1				Space 8_Time 1								
	$\Delta W: 0.039\%$		$\Delta W: 1.079\%$				$\Delta W: 4.304\%$								
	$W = 1.0738$		$W = 1.0842$				$W = 1.11645$								
		Basics 100%	Manuf.	Basic		Manuf.		Basic				Manuf.			
				Alim	Fuel	Oils 100%	Tech	Alim.	Fuel	Oils		Tech			
								Food	Bev	Mat	Min 100%	Ani	Chm	Mch	Man
		1.0738	1.0734	1.0482	1.0831	1.0842	1.0723	1.0446	1.0947	1.0651	1.11645	1.0123	1.0848	1.0819	1.0563
	Space 2_Time 2		Space 4_Time 2				Space 8_Time 2								
	$\Delta W: 0.114\%$		$\Delta W: 1.12\%$				$\Delta W: 4.304\%$								
	$W = 1.07455$		$W = 1.08416$				$W = 1.11645$								
		Basic 55%	Manuf. 45%	Basic		Manuf.		Basic				Manuf.			
				Alim	Fuel 33.8%	Oils 66.2%	Tech	Alim.	Fuel	Oils		Tech			
								Food	Bev	Mat	Min	Ani	Chm	Mch	Man
	Basic favorable 42%	1.182	1.07	1.11	1.200	1.093	1.068	1.109	1.127	1.114	1.34	0.975	1.094	1.068	1.065
	Manuf. favorable 58%	1.003	1.076	1.006	1.006	1.078	1.076	1.001	1.072	1.032	0.98	1.04	1.078	1.092	1.05
	Space 2_Time 4		Space 4_Time 4				Space 8_Time 4								
	$\Delta W: 0.109\%$		$\Delta W: 1.125\%$				$\Delta W: 4.304\%$								
	$W = 1.07450$		$W = 1.0847$				$W = 1.1164$								
		Basic 55.4%	Manuf. 44.5%	Basic		Manuf.		Basic				Manuf.			
				Alim.	Fuel 34%	Oils 66%	Tech	Alim.	Fuel	Oils		Tech			
							Food	Bev	Mat	Min 100%	Ani	Chm	Mch	Man	
Alim. 14%	1.03	0.97	1.19	0.97	0.99	0.97	1.20	1.12	1.02	0.97	0.96	0.99	0.96	0.98	
Fuel 25%	1.21	1.08	1.07	1.26	1.1	1.08	1.06	1.13	1.12	1.48	0.96	1.10	1.08	1.06	
Oil 36%	1.04	1.07	0.99	1.06	1.13	1.07	0.99	1.04	1.08	1.03	1.06	1.13	1.07	1.06	
Manu 25%	1.03	1.13	1.04	1.02	1.05	1.14	1.03	1.13	1.02	1.02	1.04	1.05	1.17	1.09	
Space 2_Time 8		Space 4_Time 8				Space 8_Time 8									
$\Delta W: 0.09\%$		$\Delta W: 1.113\%$				$\Delta W: 4.461\%$									
$W = 1.07431$		$W = 1.0845$				$W = 1.118$									
	Basic 56.9%	Manuf. 43.1%	Basic		Manuf.		Basic				Manuf.				
			Alim.	Fuel	Oils	Tech	Alim.	Fuel	Oils		Tech				
							Food	Bev 20%	Mat	Min 80%	Ani	Chm	Mch	Man	
Food 6%	1.08	0.92	1.26	1.01	0.95	0.92	1.29	1.13	1.01	1.00	0.99	0.95	0.91	0.94	
Bev 19%	0.99	1.03	1.06	0.97	1.06	1.03	1.03	1.37	1.01	0.91	0.89	1.06	1.04	1.02	
Mat 6%	1.01	1.02	1.00	1.02	1.06	1.02	1.03	0.88	1.09	0.93	0.96	1.06	1.02	1.02	
Min 36%	1.18	1.10	1.08	1.22	1.11	1.10	1.08	1.07	1.10	1.43	0.99	1.11	1.11	1.08	
Ani 17%	1.01	1.09	1.00	1.02	1.09	1.09	1.00	1.01	1.06	0.99	1.32	1.09	1.11	1.07	
Chm 11%	0.98	1.05	0.90	1.02	1.10	1.05	0.89	0.98	1.04	0.98	0.93	1.10	1.06	1.02	
Mch 3%	1.07	1.18	1.09	1.04	1.01	1.19	1.09	1.10	1.02	1.10	0.84	1.01	1.21	1.15	
Man 3%	1.23	1.30	1.18	1.23	1.27	1.30	1.19	1.16	1.22	1.25	1.23	1.27	1.30	1.30	

FIGURE 5: Space-time fitness matrixes of Japanese exports 1974–2010. Fitness values are calculated as geometric means based on annual averages. Horizontal types are based on Figure 4 (see for full names of types). Vertical environmental states are identified by superior relative fitness of a type within a squared split (see text). Fitness optimization suggests placing resources only at the shaded types, while nothing should be bet on the nonshaded ones. Optimal population shares were calculated with constrained nonlinear optimization algorithms in R (NLOptr). The values we calculated with the Augmented Lagrangian Algorithm, using the Broyden-Fletcher-Goldfarb-Shanno method (auglag-lbfsg), and confirmed with the Sequential Quadratic Programming method (auglag-slsqp), as well as with Sequential Quadratic Programming (slsqp).

fine-grained distinction would add some 1.2% of additional fitness to the population growth rate.

3.3. *The Fitness Gains of Fine-Graining.* Figure 7 visualizes the proportionality relation of complementary variety from (7) for the different cases of Japanese exports and migrating

birds in North America. Each point refers to a specific coarse-graining and a specific strategic intervention. The larger the complementary variety, the larger the fitness value of strategic cooperation among types stemming from fine-graining.

Figure 8 visualizes the resulting fitness gains of strategic cooperation according to increasingly detailed population

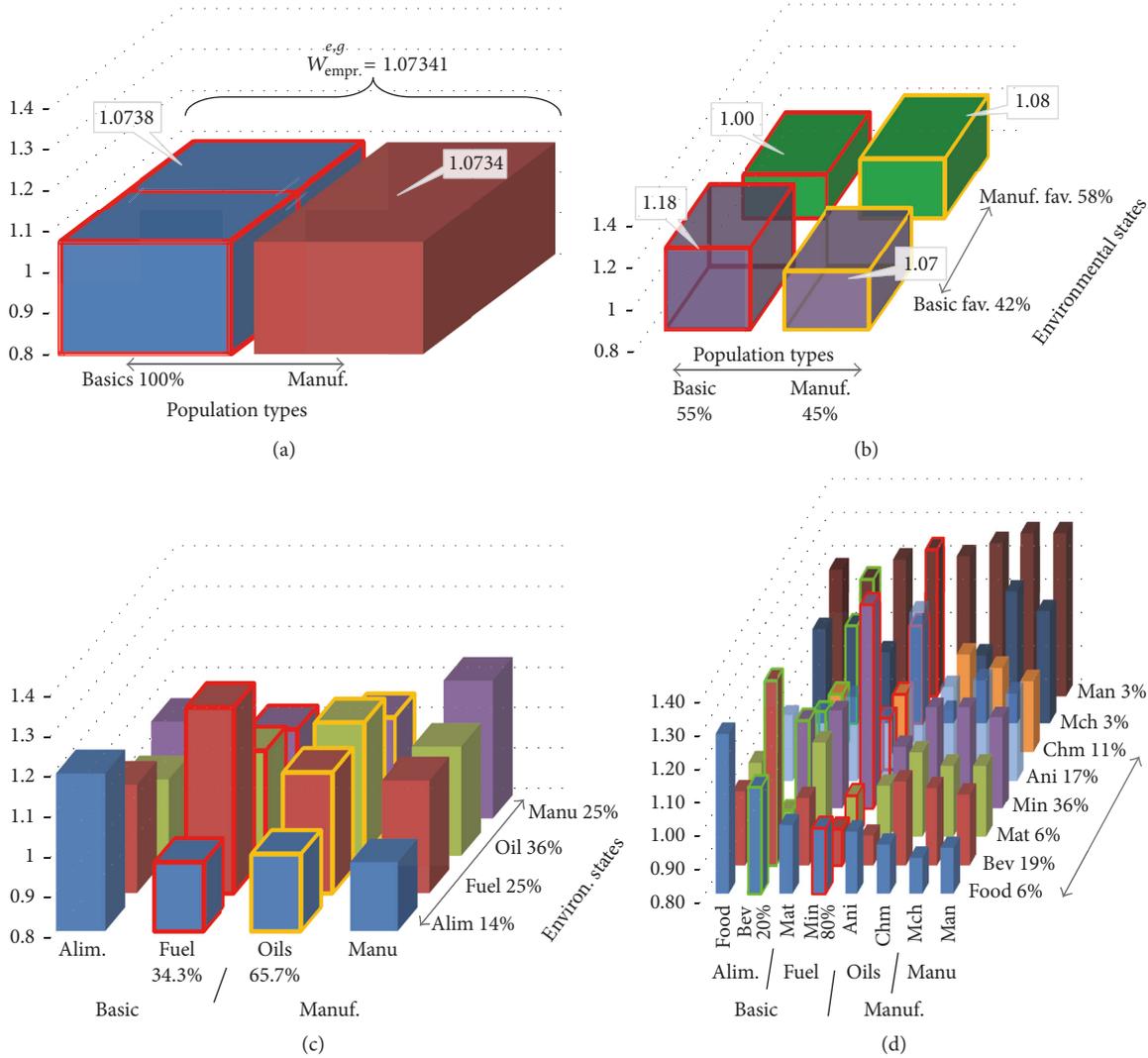


FIGURE 6: Fine-graining of the Japanese export economy (1974–2010), visualization of selected cases from Figure 5: (a) Space 2_Time 1; (b) Space 2_Time 2; (c) Space 4_Time 4; (d) Space 8_Time 8. Optimization suggests only working with the types that are outlined with colored borders, while neglecting the other types.

types (horizontal x -axis) and environmental states (vertical y -axis) (i.e., Figure 8(a) visualizes the values reported in Figure 5). It shows that more detailed distinctions among types tend to lead to higher fitness gains than more detailed distinctions among environmental states. It turns out that, in both cases, the distinction among more environmental states than the number of population types actually slightly decreased fitness: the highest fitness increase was detected with as many types as states (a squared fitness matrix, as depicted in Table 1). This is not necessarily the case for all fitness landscapes, as sometimes the distinctions among more environmental states than types can lead to an increase in population fitness (for a proof, see Supporting Information S3.3). However, this presented mathematical counterexample might also be a theoretical exception (“angels dancing on the head of a pin”), while our empirical data suggests that a match between the number of states and types plays an important role in the exploitation of fitness landscapes. Squared fitness

matrixes with one specialized type per state are often the starting assumption in the bet-hedging literature [2, 14, 15, 50], which brings us to the final discussion that links our results to this well-known literature.

4. Discussion: Optimal Intervention

The main contribution of this paper is the derivation of conditions for cooperative resource redistribution in terms of information theoretical measures. The derived proportionality relation of complementary variety applies for any strategic intervention that holds the population shares constant in an uncertain, fluctuating environment. Holding the shares constant is the best intervention one can do under “uncertainty” regarding the realization of the next environment (if the future environmental state is known, there is no uncertainty and one can do better). For our empirical exploration we used a series of numeric optimization algorithms to detect the

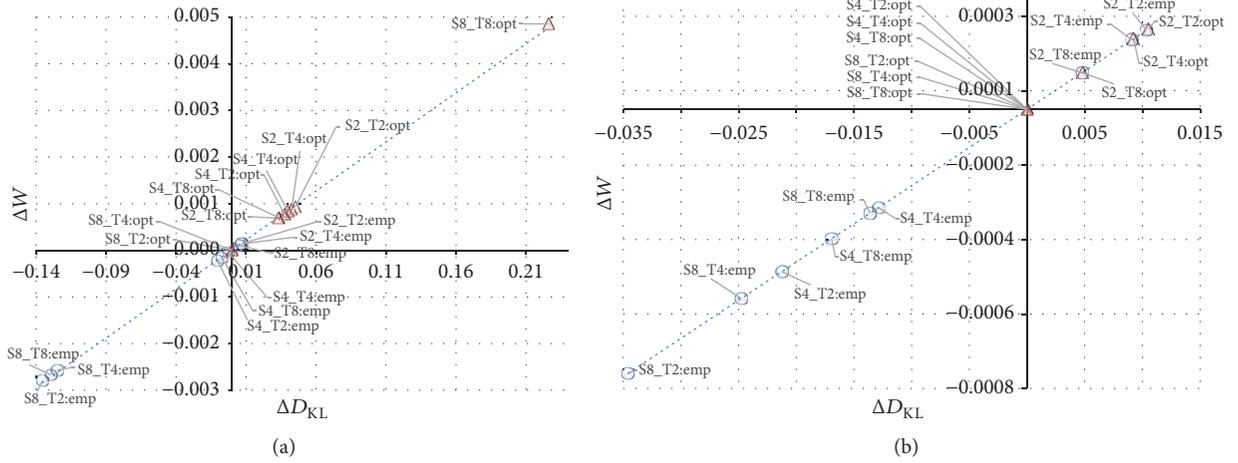


FIGURE 7: Empirical application of the proportionality of complementary variety of (7): $\Delta W \propto \Delta D_{KL}$. For the cases of (a) Japanese exports and (b) migrating birds in North America. The deltas are calculated for the empirical case of natural selection versus the case that the strategic cooperation holds the initial empirical population distribution constant over all periods (S_T:emp); and for the case that compares the population distribution that turns out to be optimal with the case in which this same optimized distribution would be the starting distribution of competitive selection over the entire period (S_T:opt). Optimal population shares were calculated with constrained nonlinear optimization algorithms in R (NLOptr).

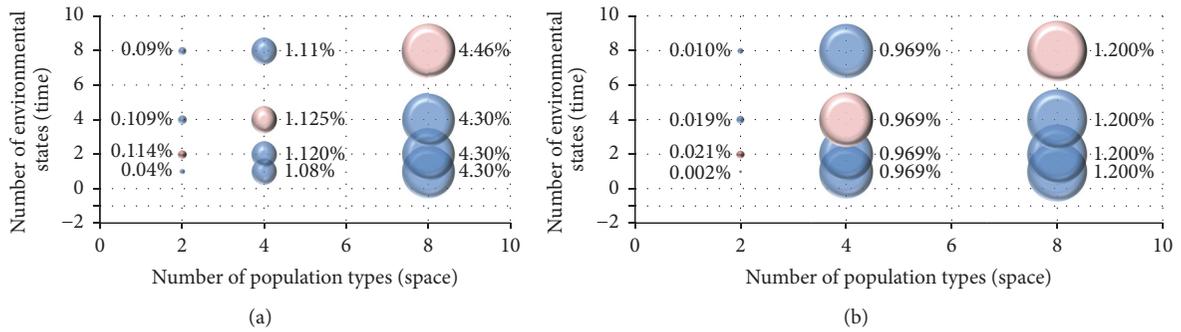


FIGURE 8: Visualization of the fitness gain of fine-graining $\Delta W = W_{SC}^{e,g} - W_{CS(emp)}^{e,g}$. (a) Japanese exports (as reported in Figure 5); (b) North American population of migrating birds. Size of bubbles refers to the fitness gain, while the largest ΔW per population split (vertical alignment) is marked reddish.

optimal portfolio distribution that is to be held constant (see Figure 5, i.e., Broyden-Fletcher-Goldfarb-Shanno, Sequential Quadratic Programming, and Sequential Quadratic Programming). There is also a well-known analytical result that allows determining the optimal distribution (for all practical purposes, at least for binary cases with two types). The analytical setup is given by the literature of bet-hedging, which goes at least back to ideas from Bernoulli in 1738 [51]. Bet-hedging is a special case of strategic cooperation; it is the optimal strategic cooperation. Being a special case, our results about complementary variety also provide the condition *sine qua non* for bet-hedging.

In his pioneering work, Kelly [2] has shown that if there is only one type that survives per environmental state (perfectly specialized type, or “winner takes it all”), then the optimal correspondence between $P(E)$ and $P(G)$ is one-to-one proportional. So if a certain environmental state shows up

$x\%$ of the time, the population should always allocate exactly $x\%$ of its resources on the type that is specialized for this state. With a diagonal fitness matrix, in order to survive, types are forced to complement the fitness strengths and weaknesses of each. Without cooperation, nobody will survive in the long term, since sooner or later an environmental state will come along that is fatal for each type. A diagonal fitness matrix exhibits an extreme case of complementary variety.

In cases where the fitness matrix is not diagonal (more than one nonzero value per row and/or column), bet-hedging optimization leads to the suggestion to distort proportionality between $P(E)$ and $P(G)$ by the mixed fitness matrix [9, 10, 15, 16]. In some constellations, it turns out to be beneficial not to allocate any resources to some types (referred to as being outside the “region of bet-hedging” [10, 14–16]). These types are superfluous (see, e.g., the cases of Space 4 and Space 8 in Figure 5).

Employing a “clutter” of matrix algebra (“a clutter is the universal measure of just-enough yuckiness to deter referees and editors from reading proofs of lemmas” (Weitzman, 1998; p. 338)) it can be shown that if the entries of the fitness matrix are linearly dependent among different environmental states, then bet-hedging over all possible types is never possible (i.e., is always outside the so-called “region of bet-hedging”) (for the formal proof see Supporting Information S4). Linear dependence of the matrix means that the fitness values of at least two environmental states are perfectly positively correlated. This suggests that the distinction among these different environmental states is redundant. On the other extreme, Supporting Information S4 also shows that the only way bet-hedging among all types is always possible is if the underlying fitness matrix is diagonal (exactly one nonzero entry in each row and each column and 0s elsewhere). In a diagonal fitness matrix the rows (environmental states) are perfectly negatively correlated and assured to be independent. This is the groundbreaking case that Kelly studied [2] and again the most extreme case of complementary variety.

In between these two extremes, the size of the region of bet-hedging—and therefore the likelihood of being able to optimize a strategic intervention—is defined by the degree of specialization among types. This refers to the difference between type that has the highest fitness in a particular environmental state (most specialized) and the less fit types in this state. The larger this difference, the larger the area of bet-hedging (this can be shown with the matrix determinant; for details see Supporting Information S4). Specialization of different types to respective environmental states implies complementary variety, as the different types potentially complement each other in different environments. This implies that the higher the degree of specialization, the larger the complementary variety, the larger the area of bet-hedging, and therefore the larger the chance of optimizing fitness with the help of all different types.

5. Conclusions

The decision if it is beneficial to intervene and to cooperate in evolutionary dynamics in uncertain, stochastic environments hinges on the shape of the underlying fitness landscape, including the information we have about it. We derived a measure that allows us to quantify the potential benefit of strategic intervention. Its interpretation leads to the concept of complementary variety of different population types in different environmental states. If there is complementary variety, it is beneficial for types to cooperate in order to exploit this complementarity. The more type fitness is skewed to opposing directions in different environmental states, the more the specialization of different types to different environmental states is extreme, the larger the complementary variety among them is, and the proportionally larger the potential benefit of strategic cooperation over competitive selection is. An important special case is Kelly’s bet-hedging with one perfectly specialized type per state [2]. In this case, types need to collaborate and cross-subsidize each other in order to survive.

Complementary variety is the necessary condition for strategic intervention to trump the “blind forces” of natural selection. In this sense, the identified proportionality relation also provides a measure for the acceptable cost of any strategic intervention. This cost has two components. For one, it is necessary to learn the environmental distribution and the related fitness values. Furthermore, resource redistribution might be costly. In many cases, there might be a diverse range of costs for redistribution among different types and levels. In reality, redistribution might not be possible at all. In practice, it might often be the case that more detailed lower levels provide both more reliable information and more easily accessible resource redistribution possibilities, while dynamics on higher levels might be more obscure and make it more difficult of exchange resources. This would suggest that lower levels lend themselves more to strategic cooperation, while higher levels would lend themselves more to competitive selection, but this hypothesis is to be empirically tested. This discussion about information and evolution, combined with the information theoretic logic presented here, provides multiple connection to the quickly growing body of literature that links evolutionary dynamics and information theory [11, 14–18, 30, 39, 40, 43–45].

Another major limitation to our results is that we assume the fitness matrix to be fixed and our time series stationary. This assumes that intervention would not alter the fitness matrix and neglects density dependence. Allocating all 2 million birds along the Northern border of the USA (see Figure 4) might be optimal if the identified fitness matrix is fixed, but in reality this overpopulation of migrating birds might saturate the carrying capacity of this region. Density dependence continues to be an active field of research for both naturalists and social scientists, but general conclusions seem difficult to obtain given the vast amount of possible interaction effects [17, 52, 53].

Finally, in our empirical exploration we have seen that our concept of complementary variety lends itself as criteria for the creation of a classification system of population types and environmental states. We have used a very simplistic way of doing so, even when adopting previously existing type taxonomies. There are surely ways to optimize the classification of types and states to maximally benefit from complementary variety. An ingenious classification system could even create different levels of complementary variety over multiple levels. The number of possible multilevel hierarchies to partition the total population is very large. It is defined by the so-called Bell number. A population with only 53 individuals allows for more multilevel partitions than there are atoms on earth (some 10^{50}), while 75 individuals allow for more different multilevel hierarchies than the estimated number of atoms in the observable universe (some 10^{80}). In a biological or social behavioral sense, certainly not all possible partitions are meaningful or allow for resource exchange. But even with these constraints we are still left with a myriad of combinatorial possibilities. The concept of complementary variety can help in the search for classification systems that identify meaningful types and environmental states over multiple levels.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

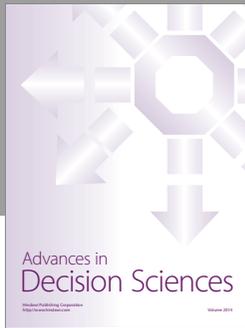
Acknowledgments

The author is indebted to Steve Frank and several critical reviewers for continuously encouraging him to look deeper, to Jim Crutchfield for continuously exposing him to the depth of information theory, and to Matina Donaldson-Matasci, Michael Lachmann, Sarah Marzen, David Wolpert, Gerhard Kramer, Nihat Ay, Poong Oh, Peter Monge, George Barnett, and other faculty members of the Santa Fe Institute for direct or indirect discussions resulting in this paper.

References

- [1] C. Darwin, *On the Origin of Species by Means of Natural Selection; Or, The Preservation of Favoured Races in the Struggle for Life*, 1861.
- [2] J. Kelly, "A new interpretation of information rate," *The Bell System Technical Journal*, vol. 35, pp. 917–926, 1956.
- [3] H. A. Latané, "Criteria for choice among risky ventures," *Journal of Political Economy*, vol. 67, no. 2, pp. 144–155, 1959.
- [4] P. H. Algoet and T. M. Cover, "Asymptotic optimality and asymptotic equipartition properties of log-optimum investment," *The Annals of Probability*, vol. 16, no. 2, pp. 876–898, 1988.
- [5] L. E. Blume and D. Easley, "Economic natural selection," *Economics Letters*, vol. 42, no. 2–3, pp. 281–289, 1993.
- [6] Keynes, *The General Theory of Employment Interest and Money*, Palgrave Macmillan, London, UK, 1936.
- [7] D. J. Wilson, "Fiscal spending jobs multipliers: evidence from the 2009 American recovery and reinvestment act," *American Economic Journal: Economic Policy*, vol. 4, no. 3, pp. 251–282, 2012.
- [8] D. Cohen, "Optimizing reproduction in a randomly varying environment," *Journal of Theoretical Biology*, vol. 12, no. 1, pp. 119–129, 1966.
- [9] P. Haccou and Y. Iwasa, "Optimal mixed strategies in stochastic environments," *Theoretical Population Biology*, vol. 47, no. 2, pp. 212–243, 1995, <https://doi.org/10.1006/tpbi.1995.1009>.
- [10] M. C. Donaldson-Matasci, M. Lachmann, and C. T. Bergstrom, "Phenotypic diversity as an adaptation to environmental uncertainty," *Evolutionary Ecology Research*, vol. 10, no. 4, pp. 493–515, 2008.
- [11] E. Kussell and S. Leibler, "Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments," *Science*, vol. 309, no. 5743, Article ID 1114383, pp. 2075–2078, 2005, <https://doi.org/10.1126/science.1114383>.
- [12] M. Salathé, J. V. Cleve, M. W. Feldman, and M. Salathé, "Evolution of stochastic switching rates in asymmetric fitness landscapes," *Genetics*, vol. 182, no. 4, Article ID 103333, pp. 1159–1164, 2009, <https://doi.org/10.1534/genetics.109.103333>.
- [13] R. Levins, "Theory of fitness in a heterogeneous environment. I. the fitness set and adaptive function," *The American Naturalist*, vol. 96, no. 891, pp. 361–373, 1962.
- [14] C. T. Bergstrom and M. Lachmann, "Shannon information and biological fitness," in *Proceedings of the IEEE Workshop on Information Theory*, pp. 50–54, 2004, <http://octavia.zoology.washington.edu/publications/BergstromAndLachmann04.pdf>.
- [15] M. C. Donaldson-Matasci, C. T. Bergstrom, and M. Lachmann, "The fitness value of information," *Oikos*, vol. 119, no. 2, pp. 219–230, 2010.
- [16] O. Rivoire and S. Leibler, "The value of information for populations in varying environments," *Journal of Statistical Physics*, vol. 142, no. 6, pp. 1124–1166, 2011.
- [17] D. Cherkashin, J. D. Farmer, and S. Lloyd, "The reality game," *Journal of Economic Dynamics and Control*, vol. 33, no. 5, pp. 1091–1105, 2009.
- [18] H. H. Permuter, Y.-H. Kim, and T. Weissman, "Interpretations of directed information in portfolio theory, data compression, and hypothesis testing," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3248–3259, 2011, <https://doi.org/10.1109/TIT.2011.2136270>.
- [19] R. Levins, *Evolution in Changing Environments: Some Theoretical Explorations*, Princeton University Press, 1968.
- [20] T. Hens and K. R. Schenk-Hoppe, "Evolutionary finance: introduction to the special issue," *Journal of Mathematical Economics*, vol. 41, no. 1–2, pp. 1–5, 2005.
- [21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley-Interscience, Hoboken, NJ, USA, 2nd edition, 2006.
- [22] R. C. Feenstra, R. E. Lipsey, H. Deng, A. C. Ma, and H. Mo, "World Trade Flows," Tech. Rep. 1962–2000, National Bureau of Economic Research, 2005, <http://www.nber.org/papers/w11040>.
- [23] R. Hausmann, C. A. Hidalgo, S. Bustos et al., *The Atlas of Economic Complexity: Mapping Paths to Prosperity*, Harvard University Center for International Development, MIT Media Lab., 2011, <http://atlas.media.mit.edu/>.
- [24] UNSD, *UN COMTRADE database (International Merchandise Trade Statistics (IMTS))*, United Nations Statistics Division, New York, NY, USA, 2012, <http://comtrade.un.org/>.
- [25] K. L. Pardieck Jr., J. D. Ziolkowski Jr., and M. A. R. Hudson, *North American Breeding Bird Survey Dataset 1966–2014*, Patuxent Wildlife Research Center, 2015, <https://www.pwrc.usgs.gov/bbs/RawData/>.
- [26] L. A. Hug, B. J. Baker, K. Anantharaman et al., "A new view of the tree of life," *Nature Microbiology*, vol. 1, no. 5, Article ID 16048, 2016.
- [27] O. A. Guerrero and R. L. Axtell, "Employment Growth through Labor Flow Networks," *PLoS ONE*, vol. 8, no. 5, Article ID e60808, 2013, <https://doi.org/10.1371/journal.pone.0060808>.
- [28] M. Platt, "Animal cognition: monkey meteorology," *Current Biology*, vol. 16, no. 12, pp. R464–R466, 2006.
- [29] M. Hilbert, "Big data for development: a review of promises and challenges," *Development Policy Review*, vol. 34, no. 1, pp. 135–174, 2016.
- [30] M. Hilbert, "The more you know, the more you can grow: an information theoretic approach to growth in the information age," *Entropy*, vol. 19, no. 2, p. 82, 2017.
- [31] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
- [32] R. A. Fisher, *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford, UK, 1st edition, 1930.
- [33] W. J. Ewens, "An interpretation and proof of the fundamental theorem of natural selection," *Theoretical Population Biology. An International Journal*, vol. 36, no. 2, pp. 167–180, 1989.
- [34] S. A. Frank, "The price equation, fisher's fundamental theorem, kin selection, and causal analysis," *Evolution*, vol. 51, no. 6, pp. 1712–1729, 1997.
- [35] G. R. Price, "Fisher's 'fundamental theorem' made clear," *Annals of Human Genetics*, vol. 36, no. 2, pp. 129–140, 1972.

- [36] J. F. Crow and M. Kimura, *An introduction to population genetics theory*, Blackburn Press, Hoboken, NJ, USA, 1970.
- [37] M. Kimura, "Natural selection as the process of accumulating genetic information in adaptive evolution," *Genetical Research*, vol. 2, no. 1, pp. 127–140, 1961.
- [38] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, Article ID 1177729694, pp. 79–86, 1951, <https://doi.org/10.1214/aoms/1177729694>.
- [39] E. Akin, *The Geometry of Population Genetics*, Springer, Berlin, Germany, 1979, http://link.springer.com/chapter/10.1007/978-3-642-93128-4_1.
- [40] A. Fujiwara and S. Amari, "Gradient systems in view of information geometry," *Physica D: Nonlinear Phenomena*, vol. 80, no. 3, pp. 317–327, 1995, [https://doi.org/10.1016/0167-2789\(94\)00175-P](https://doi.org/10.1016/0167-2789(94)00175-P).
- [41] M. Harper, "The Replicator Equation as an Inference Dynamic," 2009, <https://arxiv.org/abs/0911.1763>.
- [42] Y. Sato, E. Akiyama, and J. P. Crutchfield, "Stability and diversity in collective adaptation," *Physica D: Nonlinear Phenomena*, vol. 210, no. 1-2, pp. 21–57, 2005, <https://doi.org/10.1016/j.physd.2005.06.031>.
- [43] S. A. Frank, "Natural selection. IV. The Price equation," *Journal of Evolutionary Biology*, vol. 25, no. 6, pp. 1002–1019, 2012.
- [44] S. A. Frank, "Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory," *Journal of Evolutionary Biology*, vol. 25, no. 12, pp. 2377–2396, 2012.
- [45] S. A. Frank, "Natural selection maximizes fisher information," *Journal of Evolutionary Biology*, vol. 22, no. 2, pp. 231–244, 2009.
- [46] M. Hilbert, "An information theoretic decomposition of fitness: engineering the communication channels of nature and society," *Social Science Research Network*, Article ID 2588146, 2015, <http://papers.ssrn.com/abstract=2588146>.
- [47] G. R. Price, "Selection and Covariance," *Nature*, vol. 227, no. 5257, pp. 520–521, 1970.
- [48] G. R. Price, "Extension of covariance selection mathematics," *Annals of Human Genetics*, vol. 35, no. 4, pp. 485–490, 1972.
- [49] UNSD, *National Classifications: 870 Classifications in 154 Countries*, New York, NY, USA, 2014, <http://unstats.un.org/unsd/cr/cryreg/default.asp>.
- [50] F. Mafessoni, *Limits in selection, from populations to cognition [Ph.D. thesis]*, Leipzig university, 2015, <http://pubman.mpdl.mpg.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:2271216>.
- [51] S. C. Stearns, "Daniel Bernoulli (1738): Evolution and economics under risk," *Journal of Biosciences*, vol. 25, no. 3, pp. 221–228, 2000.
- [52] C. de Mazancourt and U. Dieckmann, "Trade-Off Geometries and Frequency-Dependent Selection," *The American Naturalist*, vol. 164, no. 6, Article ID 424762, pp. 765–778, 2004, <https://doi.org/10.1086/424762>.
- [53] C. Rueffler, T. J. M. Van Dooren, and J. A. J. Metz, "Adaptive walks on changing landscapes: levin's approach extended," *Theoretical Population Biology*, vol. 65, no. 2, pp. 165–178, 2004, <https://doi.org/10.1016/j.tpb.2003.10.001>.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

