

Evolution of what?

A Network Approach for the Detection of Evolutionary Forces

Martin Hilbert ^a, Poong Oh ^b, Peter Monge ^c

^a University of California, Davis, Department of Communication
hilbert[at]ucdavis.edu; Kerr Hall 369, Davis, CA, 95616

^b University of Pennsylvania, Annenberg School for Communication

^c University of Southern California, Annenberg School of Communication

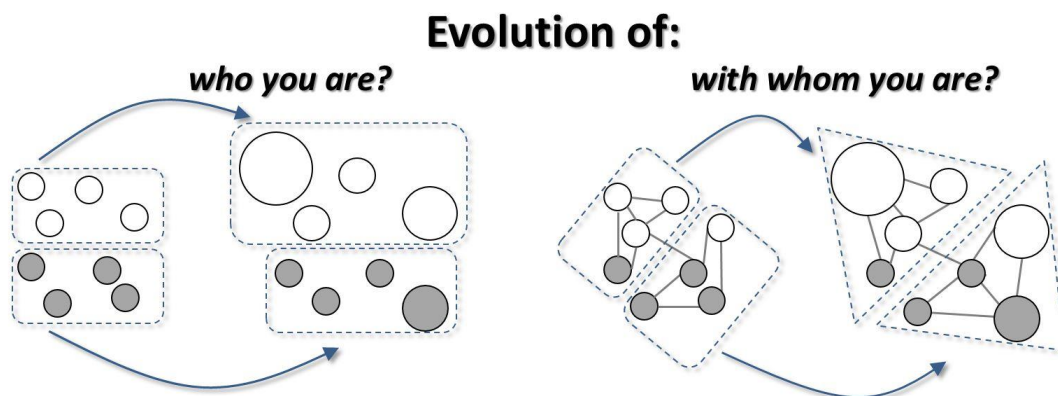
Authors' version published as: Hilbert, M., Oh, P., & Monge, P. (2016). Evolution of what? A network approach for the detection of evolutionary forces. *Social Networks*, 47, 38–46.

<http://doi.org/10.1016/j.socnet.2016.04.003>

Highlights:

- Evolutionary analysis usually focuses on traits, such as the beak length of Darwin's finches
- However, evolution acts on both who subjects are and whom they are connected with
- We analyze the effects of network metrics on traditional evolutionary equations
- We derive a formula to compare evolutionary forces for trait and network metrics
- Network metrics identify stronger evolutionary natural selection than trait partitions

Graphical Abstract:



Abstract

Structures of evolving populations are traditionally derived from traits of its members. An alternative approach uses network metrics to define groups that evolve jointly. This supposes that selection acts not only on who members are (i.e., traits) but also on to whom they are connected (i.e., interdependent relationships). This paper presents a method to meaningfully quantify differences in evolutionary forces over multiple levels of population taxonomies and tests almost 1,000 multilevel partitions of 8 empirical networked populations evolving over time. It shows that multilevel network metrics as selection criteria identifies stronger evolutionary natural selection than trait based population taxonomies.

Keywords: network evolution, network dynamics, natural selection, network partition, population taxonomy.

Contents

Introduction	3
Research Question	4
Methods.....	7
Materials and Data.....	10
Results.....	11
Conclusions	15
References	17
Supporting Information	20
(A) Notation	20
(B) Multilevel application.....	21
(C) Derivation of the multilevel decomposition	22
(D) Proof of applicability of comparative method	26
(E) Summary of datasets.....	26
References	31

Introduction

“Strictly speaking, neither genes, nor cells, nor organisms, nor ideas evolve. Only populations can evolve” (Nowak, 2006, p. 14). The concept of a population implicitly implies a criterion to separate the whole into exclusive and exhaustive parts. In evolutionary theory, these parts are often referred to as *species* or ‘*types*’. Given the difference in growth rate among different types (i.e., the types’ *fitness*) some increase and others decrease in share, changing the population constitution in relative terms. As a result, the population *evolves* over time.

Biologists and other taxonomists have long used traits, characteristics, or properties of individuals to define types. For example, Darwin famously used beak length to classify finches. In biology, fields such as phylogenetics, cladistics, and systematics have taken different and sometimes contradictory approaches to the question of what constitutes those shared traits that define a type, including morphological, physiological, molecular, behavioral, ecological, and geographic characteristics. Much more in question are meaningful population definitions in the social sciences. For example, some scholars have suggested that it is important to differentiate among different organizational forms (Hannan and Freeman, 1977; Aldrich and Ruef, 2006), others to generate a taxonomy to classify human routines (Nelson and Winter, 1985), and still others to classify social dynamics according to cultural norms and institutionalized habits (Boyd and Richerson, 2005). Unlike the beak length of finches, however, the traits of social populations—organizational forms, routines, cultures, etc.—are often not immediately quantifiable. Furthermore, there has been neither agreement nor systematic investigation about which traits should be considered. Indeed, many “official” taxonomies of social populations are rather the result of a political compromise in statistical committees than the result of a substantive discussion. For example, the United Nations Statistical Commission reports working on the basis of 870 different national classifications from 154 countries (UNSD, 2014). As a result, the evolving population of the global economy is classified into different co-evolving types not necessarily because of any substantive argument, but rather because of a compromise among different classification systems.

It is important to note that the interpretation of evolutionary dynamics of a population heavily depends on the choice of traits for defining types, and more generally, on how to partition a population into distinctive parts. To illustrate, consider the evolution of the global export economy, where countries are classified into three types—the United States, BRIC (Brazil, Russia, India, and China), and the rest of the world (Figure 1). The share of the BRIC countries has grown faster than the other two types, meaning that the BRIC type is fitter than the others (BRIC has tripled, while the other two have grown by a factor of 1.5). This has evolved the constitution of the global export economy. Further, it suggests the outstanding economic growth of the BRIC type.

What if the countries were classified in a different way? An easy answer is that the constitution would have changed differently and suggested a different interpretation, although the overall change of the global total remained the same. In short, the classification criteria that defines the *types* of any social population also defines the kinds of evolutionary dynamics that can be detected and interpreted. As such, the definition of types has traditionally been subject to “context, intellectual goal and subjective bias about what is ultimately meaningful” (Frank, 2012a, p. 230). Then, the question becomes: What classification criterion should be adopted, that is, how should a social population should be grouped in order to detect and analyze the evolutionary dynamics in meaningful ways?

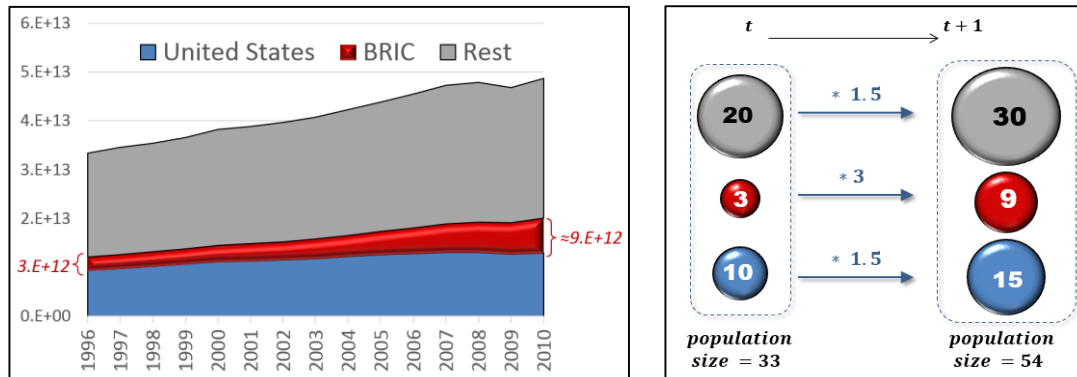


Figure 1: Schematization of evolutionary selection among three types of the global export economy (a) historical data for U.S., BRIC (Brazil, Russia, India, and China), and Rest of the world (of a total of 118 countries); (b) schematization among three corresponding illustrative types (for more on the data, see Supporting Information).

Adding a new perspective to this longstanding debate, Monge, Heiss, and Margolin (2008) have suggested that it would be beneficial to switch focus from traits of types to their network properties as the basis for defining evolving populations: “Traditionally, evolutionary theorists and researchers have examined organizational populations that comprise organizational communities by focusing on their properties rather than on the networks that link them. However, a full understanding of the evolution of organizational communities requires insight into both organizations and their networks.” (p. 449). Studies have started to explore the evolutionary role of network structures in biology (e.g. Kossinets and Watts, 2006; Palla, Barabási, and Vicsek, 2007; Cantor et al., 2015) and in social evolution (e.g. Hausmann et al. 2011; Lee and Monge, 2011). This implies that evolutionary selection does not only favor or disfavor certain population types because of who they are, but also because of who they are with. In terms of Darwin’s finches, some finches, despite having the same beak length as others, could outbreed because of the relationships in their network. Similarly, could it be that some groups of countries achieve higher fitness in the context of global economy, not just because of their traits like the number of inhabitants or income level, but because of how they connect with others? This article asks if there are differences in the strength of evolutionary forces between when a population is classified according to some common trait and according to network metrics.

Research Question

Increasingly, longitudinal network data have become more available, which has led to different research venues related to dynamics in network analysis (Stokman and Doreian, 1997; 2001; Snijders, 2001). Much attention has been paid to modelling the formation and/or resolution of network ties over time, such as in stochastic actor-oriented models (Snijders, 1996; Snijders et al. 2010), relational event frameworks (Butts, 2008) or economic incentives mechanisms (Jackson and Watts, 2002; Jackson 2010). A major goal of these approaches is to estimate the likelihood of network ties among nodes, $p(x_{ij} = 1)$, given either the endogenous or exogenous characteristics of the nodes, or both, and thereby, to identify significant characteristics that contribute to tie formation/resolution.

This study also works with longitudinal networks, but does not ask about tie formation. We explore the structure of networks to create groups of nodes that are defined by network links and then use these groups to define types. We then apply these networked types to the analysis of evolutionary dynamics in terms of changes in population fitness. We compare the network-based classification (i.e., to whom they are connected) with the traditional trait-based classification (i.e., who they are) and contrast differences in the involved evolutionary forces of change. In short, we do not ask how the network ties evolve, but—following the traditional literature of evolutionary dynamics—we ask about the fitness of the population and how different types contribute to it (given a certain network structure or node trait).

Ever since the first versions of a “mathematical theory of natural and artificial selection” (Haldane 1924; 1934), the formal evolutionary analysis in biology, ecology, anthropology, economics and other social sciences has defined fitness in terms of the growth factor of the population. Biologists often refer to it as “the number of offsprings” or the “rate of reproduction”, while game theorists refer to it as a “payoff”, and economists as a “return on investment” (e.g., Nowak, 2006, pp. 14 & 55). We follow this “growth factor” interpretation of fitness and use \bar{W} to denote the fitness of the entire population. It is defined as the number of offspring at the time $t + 1$ divided by the number of ancestors at time t , $\bar{W} = \left[\frac{\text{units}^{t+1}}{\text{units}^t} \right]$. For example, suppose that a population with 8 units at time t grows in size and comes to have 18 offspring units at time $t + 1$ (Figure 2). Then, the overall fitness \bar{W} is equal to 2.25 ($= 18/8$).

What drives the evolution of the population constitution are differences in type fitness. Types with superior relative fitness gain population share, and will dominate the population in the long run, while types with inferior fitness will be selected against and decline over time. The fitness of each type is denoted by w . For example, the type with the trait ‘white’ in Figure 2a has a type fitness of 3 (growing from a total value of 4 to 12), while the shaded type has a fitness of 1.5. selection favored white and as a result the new population is two thirds white. That is, the group of white units multiplies by a growth factor of 3 on average, while the group of shaded units multiplies by an average fitness of 1.5. A type with a fitness of $w = 0$ would die out and go extinct.

Based on this common logic, the different versions of Figure 2 visualize the basic idea behind this article. All four images show the same population, but the population is classified in four different ways. The Figure visualizes that different ways of classifying populations also changes the respective type fitness w . This is what leads to different strengths in the underlying evolutionary selection pressures. For example, the grouping according to ‘shape’ in Figure 2b does not imply any selection, as it exhibits a uniformly distributed 50 % - 50 % population both before and after updating. Type fitness is equal to 2.25 for both types. Note that the overall fitness \bar{W} remains the same, regardless of how the population is classified. The population grew, but it did not evolve, as no selection occurred.

In Figures 2a and 2b, the traits that define types are color and shape. In Figure 2c the population is classified in terms of a network partition (grouping those who are more connected to each other and less connected to the others), and Figure 2d defines the population structure in terms of the network metric of degree centrality (grouping those with three degrees or more in one group, and those with less degrees in another). As a result of these different population classifications, the involved fitness levels of the identified types are again distinct, which affects the degree of selection.

The question of the effects of population classification on evolutionary dynamics hinges on the identification of meaningful evolutionary forces. Fisher’s fundamental theorem (Fisher, 1930) equates the

variance in type fitness with the strength of natural selection (Ewens, 1989; Fisher, 1930; Frank, 1997; Price, 1972b). The greater the observed variance in growth rate among the types, the greater the change in the population constitution (as the fitter gain in share), which in turn implies and therefore stronger selection pressures. In Figure 2, the population definition based on the trait color leads to the strongest force of natural selection among the presented cases, as it distinguishes between growing white types and rather stagnating shaded types (Figure 2a). Therefore, it can be concluded that the trait ‘shaded’ is most receptive to environmental incentives. In terms of the evolutionary interpretation of fitness it is taken as a sign that the trait ‘color’ reveals *something* about the fit of the population relative to its environment, which is materialized by its ‘fitness’. On the contrary, the distinction between round and square types in Figure 2b does not identify any distinction between type fitness. Both types have the same reproduction rate of 2.25. The trait ‘shape’ seems not to capture or reflect anything important about the environment.

Therefore, while the overall population fitness is the same in all four cases, the strength of natural selection varies depending on the definition of types and how individuals are grouped. This article provides a quantitative criterion to compare evolutionary forces on populations defined differently, and therefore, it provides a method to evaluate the effects of the choice of population structure on evolutionary forces. Eight growth variables of 5 empirically evolving populations are tested. This includes data about the evolution of news, online videos, entrepreneurial fundraising, and whole economies.

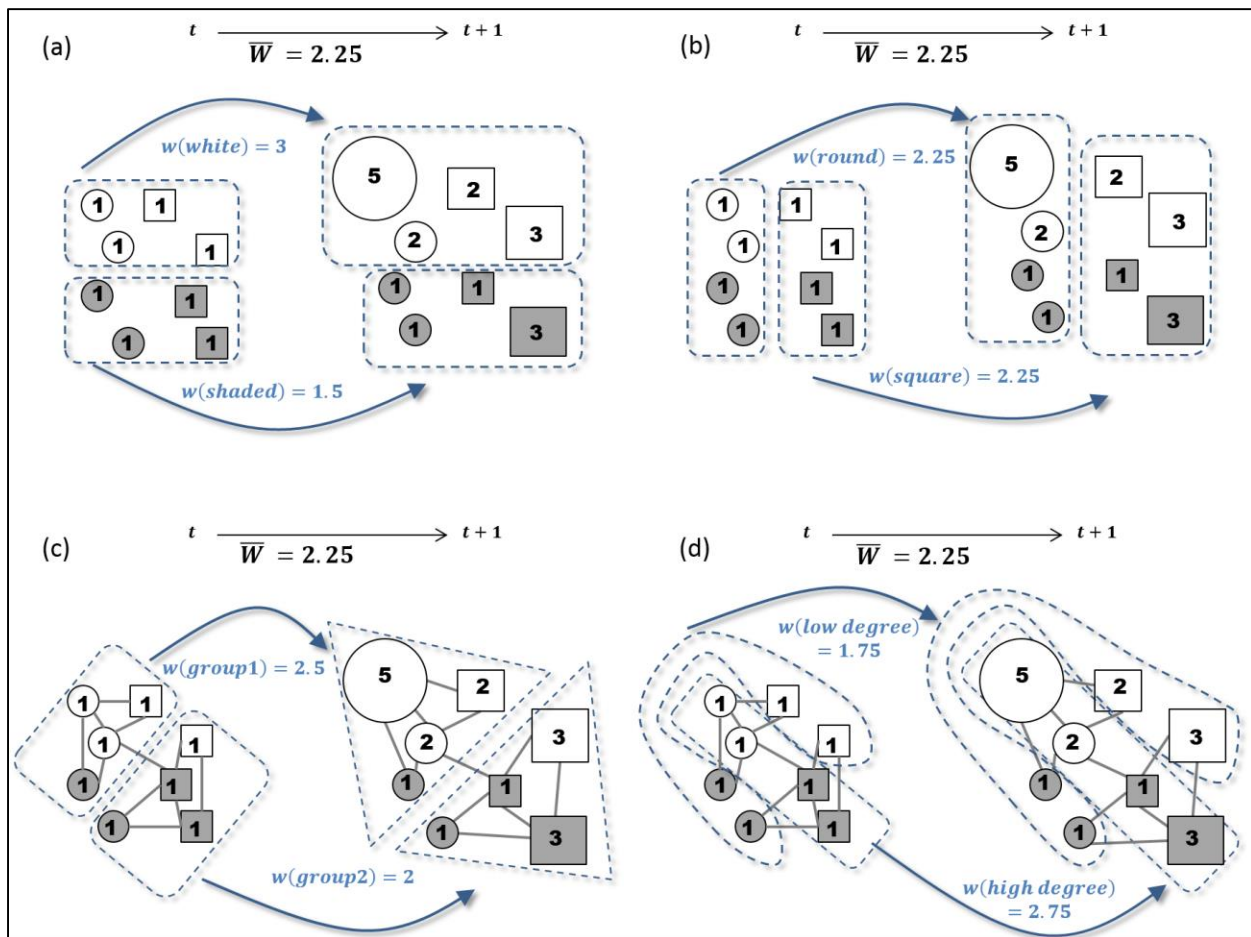


Figure 2: Schematization of alternative population definitions according to (a) the trait color; (b) the trait ‘shape’; (c) a network partition; (d) network positions. Each of the cases (a), (b), (c), (d) adds up to the same population fitness (population growth factor of $\bar{W} = \frac{18}{8} = 2.25$), but represents different evolutionary forces involved in updating. The corresponding forces of natural selection (NS), defined as the mean-normalized variance of type fitness $\frac{Var(w(g))}{\bar{W}}$, are (a) $NS(3, 1.5) = 0.25$; (b) $NS(2.25, 2.25) = 0$; (c) $NS(2.5, 2) \approx 0.03$; and (d) $NS(2.75, 1.75) \approx 0.11$.

Methods

This section presents the derivation of a formula that allows us to compare the strength of evolutionary forces for different population partitions. As Fisher himself reminded us with the first sentence of his influential 1930 book: “Natural Selection is not Evolution” (Fisher, 1930, p. vii), it is merely part of evolutionary change. A commonly used complete descriptive decomposition of all evolutionary forces has been proposed by Price (1970, 1972a). It contains natural selection as an important part, but additionally considers possible environmental changes (Frank, 1995; 1998; 2012b). Being a mathematical identity, it necessarily holds true. It decomposes the change in population fitness in terms of variances and covariances. With fitness being a growth rate, the change in fitness is essentially the acceleration / deceleration of the population dynamics, $\Delta\bar{W} = \bar{W}^{t+1} - \bar{W}^t$, where delta Δ refers to the difference by subtraction in time. The decomposition includes a multilevel recursion. We work with a modified version of the Price equation¹ (see Supporting Information for the derivation):

$$\text{change in pop. fitness} = \text{Natural Selection} + \text{Relative Acceleration} + \text{Multilevel Term}$$

$$\Delta\bar{W} = \frac{Var(w)}{\bar{W}} + Cov\left(\frac{w}{\bar{W}}, \Delta w\right) + E[\Delta w] = NS + RA + MLT \quad (1)$$

$$= \frac{Cov(w^t, w^{t+1})}{\bar{W}} + E[\Delta w] = SS + MLT. \quad (2)$$

In agreement with Fisher’s theorem, we refer to the mean-normalized variance in equation (1) as the force of natural selection (NS) (Price, 1972b; Ewens, 1989; Frank, 1997). Being a variance, it has a non-negative value, meaning that it monotonically contributes to changes in population fitness. The second term in (1) is the covariance between the relative current fitness and future change in these growth factors. It represents the change of the types’ growth factor in relation to the relative type fitness. It therefore represents the relative acceleration (RA) of types. This acceleration (or deceleration, if negative) can stem from changes in the external environment, can be endogenously produced by changes within the population, or be the result of an interaction between the types and its environment (such as in the case of density dependence with an environmental carrying capacity reference?). If current relative

¹ Replacing the freely selectable parameter z of Price’s setup with fitness w .

fitness of types ($\frac{w}{W}$) covaries positively with the change in fitness of types (Δw), it contributes positively to the increase in population fitness. Empirically, it is usually a negative counterpart to the positive contributions of natural selection (Hilbert, 2013; 2016).

The sum of these two counterbalancing evolutionary forces ($NS + RA$) adds up to the mean-normalized covariance between current and future fitness of each type (see equation (2)). This covariance is typically close to 0, implying constant population fitness on average. If negative, types that increase their share at time t decrease their share at time $t + 1$ on average, contributing to a decrease in population fitness. With ongoing deceleration, such population would quickly go extinct. If positive, the growing shares sustain their superiority in the future and therefore sustainably increase population fitness over both periods. With ongoing acceleration, such population would quickly overrun the planet. In practice, both usually balance each other out. This shows that both evolutionary forces, natural selection (NS) and relative acceleration (RA), are equally important determinants of the change in long term fitness. In the short term, one might dominate the other, but in order to achieve sustainable fitness over the long-term, both are necessary. Therefore, we call the two forces combined Sustainable Selection as represented by the following equation: $SS = NS + RA$ (see equations (1) and (2)).

We refer to the last term $E[\Delta w]$ in equations (1) and (2) as the multilevel term (MLT). It pays tribute to the fact that evolutionary dynamics consist of an “indefinite number of wavelike fluctuations which will roll on simultaneously” (Schumpeter, 1939, p. 169) on multiple levels over “hierarchic complex systems” (Simon, 1962, p. 473). The change in fitness of a certain type (Δw) can be decomposed with the same manner the change in population fitness ($\Delta \bar{W}$) was decomposed on the higher population level. What can be applied to the entire population, can also be applied within a specific type. This results in a recursive multilevel equation. The shares on lower levels are conditional on group membership on higher levels. This results in a hierarchic taxonomy. On the lowest level are indivisible units as defined by the most fine-grained level permitted by the available database (see Supporting Information for an application of such multilevel decomposition to Fig. 1).

We group the resulting terms of natural selection and relative acceleration over all levels in a multilevel taxonomy $l = [1, 2, 3, \dots, L]$, from the entire population on the highest level ($l = 1$), down to the lowest possible level of fine-graining permitted by the database ($l = L$). This results in a multilevel sum of variances from each level, NS_{Total} , and a multilevel sum of covariances from each level, RA_{Total} (see equation (3)).

$$\Delta \bar{W} = \sum_{l=1}^L NS_l + \sum_{l=1}^L RA_l + MLT_{l=L} = NS_{Total} + RA_{Total} + MLT_{l=L}. \quad (3)$$

For our comparative purposes, it is useful to realize that the multilevel sum of the variance (NS_{Total}) and covariance terms (RA_{Total}) add up to the same combined sustainable selection (SS_{Total}) effect for all different possible ways to identify type membership, as long as this multilevel logic is pursued until the lowest possible level of fine-graining (see Supporting Information). This allows for direct comparisons among quantified evolutionary forces of different structures of the same population, independent from the chosen number of levels L and from the number and size of groups on a chosen level:

$$NS_{Total} + RA_{Total} = SS_{Total} = K, \quad (4)$$

where K is constant for all possible multilevel partitions.

The number of possible multilevel hierarchies to partition the total population is very large. It is defined by the Bell number (Bell, 1938), which is the sum of all possible Stirling numbers of the second kind. For example, for a population of only 56 individuals, there are about 10^{55} possible multilevel partitions, while a population of 139 individuals can create about 10^{172} hierarchical taxonomies (Calcul, 2012). In comparison, there are estimated to be 10^{50} atoms on earth, and some 10^{80} atoms in the observable universe. Despite this unfathomable magnitude, we know that all of these possible partitions of a population add up to the same total sustainable selection effect (SS_{Total}), resulting in the same total change in population fitness. However, some ways to assign type membership result in more extreme forces of natural selection NS_{Total} and counterbalancing relative acceleration RA_{Total} , while other hierarchical taxonomies result in less extreme evolutionary forces. In terms of equation (4), the positive NS_{Total} term and the negative RA_{Total} differ in their magnitude, while their sum SS_{Total} is constant for all different ways the multilevel group membership can be constructed.

This gives us a meaningful way to compare the strength of evolutionary forces over multiple levels. We, therefore, decompose a population over different hierarchies until the lowest level and compare the strength of the evolutionary forces. We then work with the time average of the forces over a time series. For example, when working with an evolutionary trajectory over 24 points in time, we calculate equation (3) for each of the 23 sliding windows² and then build the arithmetic mean of the resulting forces NS_{Total} and RA_{Total} over the 23 periods. This gives us the time average of the population level evolutionary forces over this period.

Summing up, our method evaluates changes in fitness, which are changes in the population's growth rate. The accelerating or decelerating population dynamic is decomposed into short-term selection (which types grow faster at a certain instance?), quantified by natural selection NS , and the ongoing trend in fitness (do the fit or the unfit increase fitness?), quantified by relative acceleration RA . Together they tell us if the ones that are fit right now, will also be the fit in the future, quantified by sustainable selection SS . If we decompose a certain population classification with different groups until the lowest level of individuals, it turns out that this last quantity SS is equal for any possible multilevel taxonomy. However the two counterbalancing terms that create this total (equation (4)) can have different values. This allows us to both compare different population taxonomies (they all add up to the same total), while studying differences in the involved evolutionary forces (i.e. natural selection NS and relative acceleration RA).

We apply this method to different ways of structuring the population. Following the logic outlined in Figure 2, we test for multilevel partitions that follow different node traits (see Figure 2a and 1b); different network partitions (Figure 2c); and different network position metrics (Figure 2d). We hypothesize that network approaches to population classification result in statistically significant different evolutionary forces across all tests than when traits are used to define multilevel populations.

² The calculation of $\Delta\bar{W}$ spans 3 consecutive points in time: $\Delta\bar{W} = \overline{W^{t+1}} - \overline{W^t} = \left[\frac{\text{all units}^{t+2}}{\text{all units}^{t+1}} \right] - \left[\frac{\text{all units}^{t+1}}{\text{all units}^t} \right]$.

Materials and Data

To investigate the effects of different population classification systems, we used diverse social network data, including media-, organizational-, and country-level networks (Table 1). The current data set includes two hyperlink networks of YouTube videos—one is the network of news clips (*Democracy Now*); the other is that of scientific documentaries (*Nova PBS*)—two organizational networks from the microcredit crowdsourcing platform Kiva, and the international trade network among 118 countries (see Supporting Information for details). The number of nodes ranged from 56 to 120 (see Table 1). Each network includes evolutionary dynamics over different time periods, from 6 days to 24 years. During these periods each network was recorded at 24 points in time. Each longitudinal network was aggregated, with edges representing the average strengths of connections over the entire time period. As summarized in Table 1, the structural properties of the five networks varied: weighted network density ranged from .034 to .915, and clustering coefficient from .114 to .928. In order to test for sensitivity to the choice of the growth variable, we tested two different growth variables for the Kiva and Trade network. This means that we run our exercise twice in these cases. For example, for international trade, we analyze the evolution of the growth of trade volume in terms of the absolute amount of trade on the national levels of all 118 countries, and we also analyze it in terms of per capita trade for the 24 years between 1987 and 2010.

Network Name	Nodes	Evolutionary Growth Variable	Links	Kind of links	Number of nodes & links	Network Density	Network Clustering Coefficient	Time period	# of Traits & Network Position & Network Partitions
Democracy Now	YouTube videos	Number of views	Hyperlinks to other videos	Directed	56 & 105	0.034	0.283	12 days	10 & 10 & 2
NOVA PBS	YouTube videos	Number of views	Hyperlinks to other videos	Directed	86 & 171	0.023	0.114	12 days	10 & 11 & 2
KIVA indiv. lenders	Fundraising enterprises	Money raised: - % of credit goal - in absolute USD	Lenders in common (individuals)	Undirected	120 & 1,779	0.249	0.661	6 days	9 & 16 & 2
KIVA teams	Fundraising enterprises	Money raised: - % of credit goal - in absolute USD	Teams of lenders in common	Undirected	118 & 4,512	0.654	0.823	6 days	10 & 17 & 2
Internat. Trade	Countries	GDP in const. USD: - national total - per capita	International trade of goods	Directed	118 & 12,745	0.9153	0.9288	24 years	30 & 12/14 & 2

Table 1: Summary of empirical databases (for details and sources, see Supporting Information).

The last column in Table 1 shows the number of traits and networks metrics we used to produce different population classifications. The different traits and network metrics were chosen due to their availability, ease of access, and feasibility of creating multilevel hierarchies. Traits were specific to each case. For example, of the population of YouTube videos can be classified by the number of “Likes”

received, forming subgroups of “popular” and “unpopular” types of videos. The population of microcredit fundraising entrepreneurs can be structured according to the default rate of their partner organization (distinguishing between more and less trustworthy partnerships), among others while the evolution of international trade can distinguish between economies with fast- and slow-growing Gross Domestic Product, among others. For each case, we sorted all members according to these criteria and then grouped those that are similar.

The employed network position metrics are more similar among our cases and include the most important centrality metrics (degree, eigenvector, closeness, etc.) and other standard metrics like triad count and Burt’s (1992) constraint measure. Additionally, we execute two network partitions, namely Fiedler’s spectral partitioning method (Fiedler, 1973; see also Newman, 2010) and the Kernighan-Lin algorithm (Kernighan and Lin, 1970). These partitioning algorithms allow us to predefine group size at each level.

This is important, because if group size would differ among different cases, this change in structure can affect the variance in fitness and can, therefore, confound the magnitude of natural selection with the choice of group size. Therefore, we use a rigid rules-based approach to creating the same multilevel hierarchy for each case. We use four different rules. Rule 1 refers to a binary split on each level, resulting in a binary tree structure (Rule 1: $2 \times 2 \times 2 \times \dots$). For example, this means that the 56 nodes of the ‘Democracy Now’ network require 6 levels of bisection until the lowest level.³ Rule 2 splits the first level into 5 equal sized groups on the first level (uneven cases are decided at random), 4 on the second level, 3 on the third, etc. For example, the 120 nodes of the ‘Kiva individual lenders’ network results in exactly four levels (Rule 2: $5 \times 4 \times 3 \times 2 = 120$). Rule 3 goes the other way around from less to more groups on subsequent lower levels (Rule 3: $2 \times 3 \times 4 \dots$), and Rule 4 refers to an equal-sized quaternary tree structure (Rule 4: $4 \times 4 \times 4 \dots$). It is important to underline that this rigid rule-based population structure puts an artificial constraint on the population structure. However, it puts the same constraint on all population structures (i.e. trait-based and network-based) and allows us to focus on the strength of natural selection and relative acceleration while controlling for the potentially confounding effects of variations in group size.

Results

As a result of the combination of the different traits and network metrics with the four different hierarchies, we obtained 988 different population structures. Given equation (4), it is meaningful to directly compare the evolutionary forces of the same network. As the resulting variances vary in absolute terms among networks, for comparative reasons we rank them as equidistant percentiles from the lowest to the highest NS_{Total} and RA_{Total} , with percentiles closer to zero representing smaller forces. Figure 3 compares the average rankings, distinguishing between population structures made of (1) *trait-based*

³ $\log_2(56) \approx 5.8$. In detail, we have 2 groups on the 1st (population) level, 4 on the 2nd, 8 on the 3rd, 16 on the 4th, 32 on the 5th, and 56 nodes on the lowest (most fine-grained) 6th level.

group memberships, (2) *network position based group memberships*, and (3) *network partitions*. It shows that population classifications according to trait metrics generally result in less extreme evolutionary forces than multilevel taxonomies based on network metrics. This finding naturally accounts for both NS_{Total} (network metrics result in larger forces of Natural Selection) and its negative counterpart RA_{Total} (network metrics result in more negative Relative Acceleration). The exception to this general tendency is the Nova PBS network (see Figure 3).

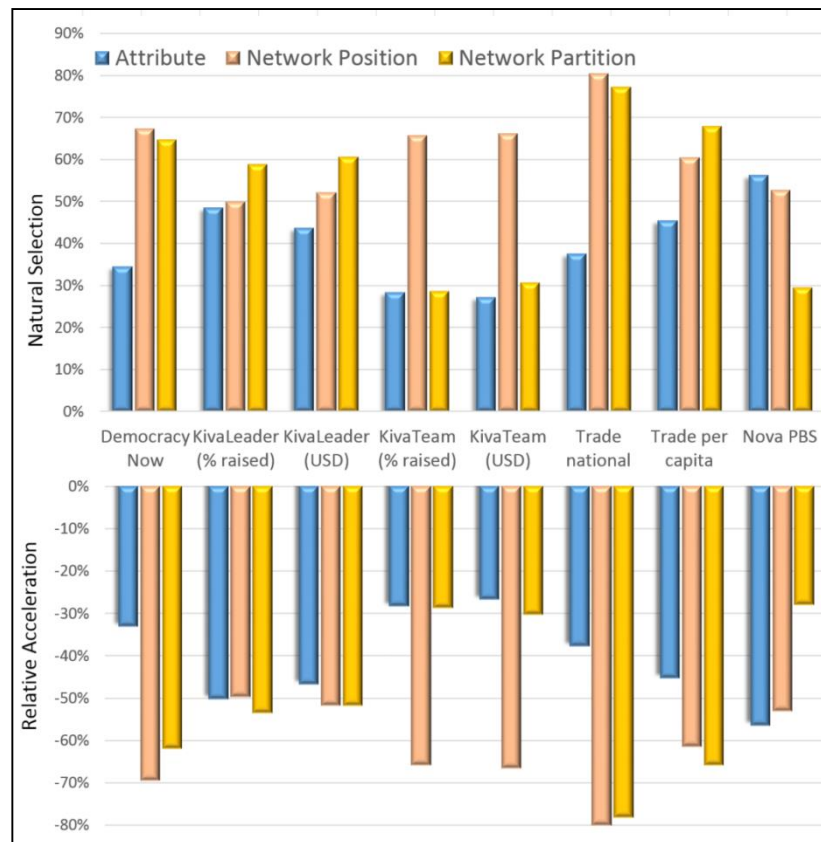


Figure 3: Average percentile rankings of evolutionary forces per multilevel population. Relative Accelerations (RA) are represented as negative percentiles as they are the negative counterpart of the positive forces of natural selection. Number of partitions: $N = 988$ consisting of DemNow (88), KivaLeader% (108), KivaLeader\$ (108), KivaTeam% (116), KivaTeam\$ (116), TradeNat (176), TradeCap (184), Nova (92).

As one of our trait metrics, we used the growth rate of the types themselves. This means that we looked at the 24 periods, and then grouped those types that grew fastest and those types that grew slowest (on average over 24 periods). If growth rates would be stationary over the 24 periods, this should clearly identify the fittest and the least fit and, therefore, lead to strong forces of selection. However, type fitness varies between the periods and, therefore, the arithmetic mean of the mean-normalized variance over 24 periods is not automatically maximized. The force of natural selection of the resulting multilevel

structure is on average within the 64 % of the strongest forces of average natural selection. This is within the range of the evolutionary force that results when creating group memberships according to network position metrics (see Figure 3). In other words, network position metrics identify fitter and less fit types on average over longer periods of time just as well as a posthoc grouping of the average fittest.

In addition to this descriptive analysis, we were also interested in whether the different trait and network partitions would be significantly different from random multilevel group memberships. This requires the creation of random partitions. Since the number of possible random partitions is very large (Bell number, see above), we first generated a total of 500,000 random partitions and then increased the sample size 20-fold to 10,000,000 random partitions. Figure 4 shows the resulting sampling distribution for NS_{Total} of the 56 members of the ‘Democracy Now’ network according to division hierarchy Rule 1 (binary split over 6 levels). The resulting distribution is not intuitive, but given that both distributions appear almost identical (see Figure 4), we consider a sampling size of 10,000,000 to be sufficient (equivalent results are obtained for all other networks).

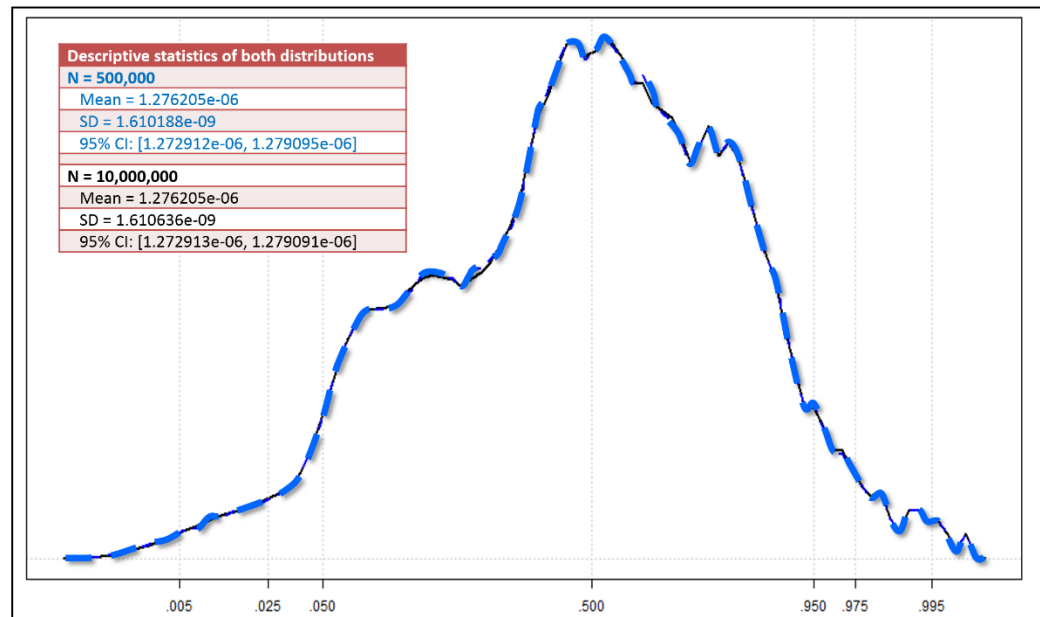


Figure 4: Sampling distribution of NS_{Total} of the 56 members of the ‘Democracy Now’ network according to hierarchy Rule 1 (binary split over 6 levels) for 50,000 (blue dashed line) and 10,000,000 (black solid line) random partitions.

Now, we checked if the NS_{Total} values obtained from our non-random trait and network partitioning were statistically significant from this random partitioning. Testing for significance at $\alpha = .05$ (two-tailed), some 16.8 % of all ‘trait-based types’ are significantly different from random, 19.7 % of all ‘network position based types’, and 23.5 % of the ‘network partitions’. It is important to notice that in this two-tailed test NS_{Total} values can be significantly higher or significantly lower than random partitions. Figure 5 shows a striking pattern with regard to the significance on the high and low end. Among all ‘trait based types’ that show a NS_{Total} significantly different from random, more have significantly smaller NS_{Total} , while the pattern turns around for network metrics. This reconfirms our previous finding that

network metrics tend to identify population structures that exhibit a relatively high force of natural selection. This accounts especially for network position metrics. Some 18.4 % of the population structures based on network position metrics show NS_{Total} that are significantly higher than any random type membership, while only 1.3 % of them are significantly lower than random partitioning.

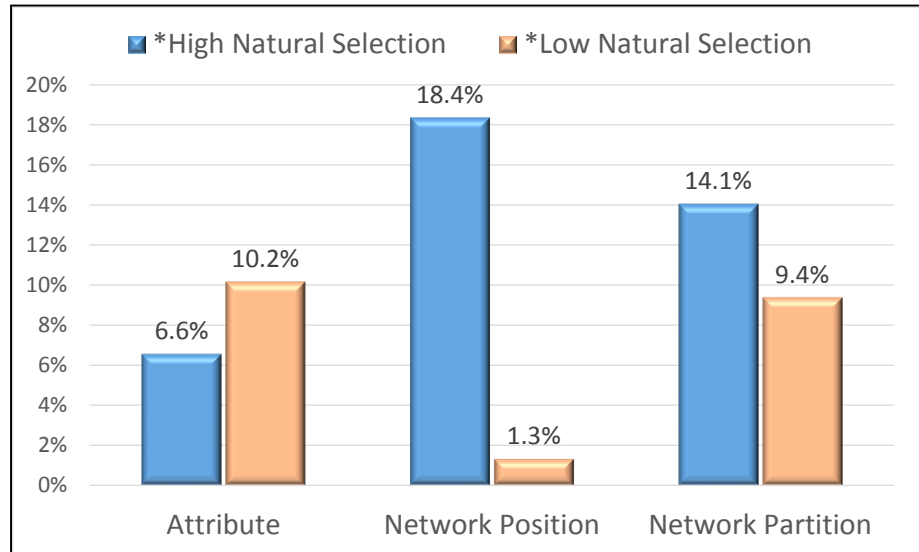


Figure 5: Percentage of population structures with NS_{Total} significantly different from random (at alpha .05, two-tailed) of all partitions of that sort. Total number of partitions: N=988 consisting of 472 trait based; 452 network position based, and 64 network partitions.

Finally, we ask if the difference between the NS_{Total} of a pair of two structures is significantly different from random (also from 10,000,000 simulations). This tests whether or not a partitioning method significantly reduces or increases NS_{Total} compared to another. We randomly choose 31,832 pairs from the possible 61,552 pairs and test against their respective sampling distributions. The result is in line with our previous finding. Figure 6 shows that in more than 20 % of the comparisons between ‘trait based types’ and a ‘network position based types’, the network metric produces a NS_{Total} that is significantly higher than the ‘trait based’ approach. The contrary is true only 3.4 % of the time. Something similar holds for ‘network partitions’. Comparing ‘network partition’ and ‘network position’ metrics (see schematizations in Figure 2c and 2d), it turns out that ‘network position’ metrics lead more than 4 times more often to significantly higher multilevel forces of natural selection.

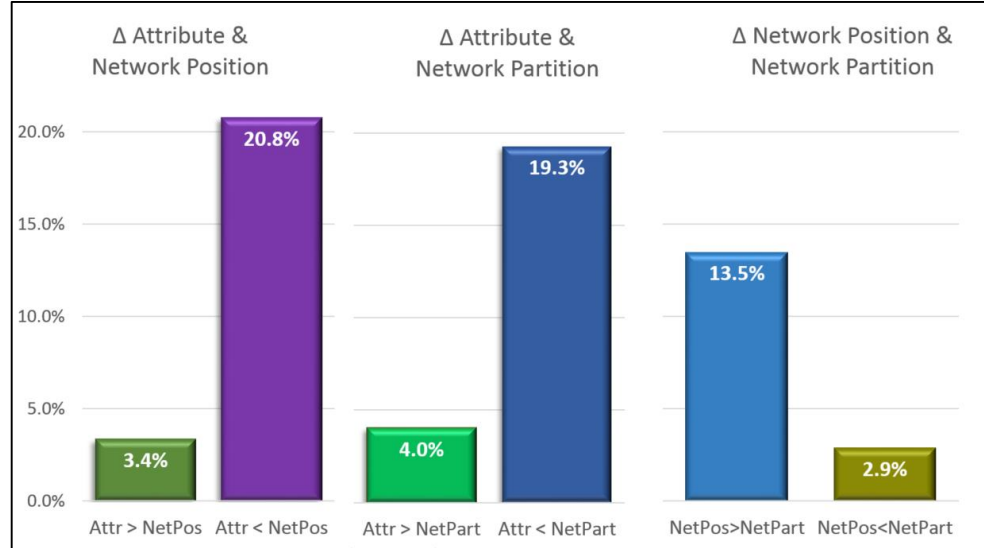


Figure 6: Pairwise differences of NS_{Total} that are significantly different from random (at the 0.1 level) as percentage of all pairwise comparisons of the indicated sort.

Conclusions

We linked network structure of the populations in question and evolutionary dynamics of those populations and presented a method to meaningfully compare the magnitude of evolutionary forces over multiple levels. We found that evolutionary forces were more extreme when we structured a multilevel population according to network metrics in comparison to using traits of the population members as criteria. This implies that network metrics are well suited to identify who of the population members are fitter and who are less fit (over changing environmental fitness landscapes). This provides a justification for using network metrics as a basis for the grouping of individuals into types in an evolving population. This perspective on population dynamics suggests that selection might not merely act on *who* individuals are, but also *with whom* they are connected. Instead of selecting for a trait, selective pressures acts on groups of individuals that occupy a certain position in the network of the population relative to other members of the population (network position based). Selection also seems to act on groups of individuals that are linked to each other (network partitions).

The reasons behind these findings are not well understood. It is likely that the reason for the strength of evolutionary forces is different when network position metrics are used and when network partitions are used. Network position metrics identify nodes that play a special role in the network structure. These nodes are not necessarily closely connected to each other. Grouping them together often joins nodes from different parts of the network, but with similar network characteristics. In contrast, network partitions join a group of nodes that might be rather diverse, but more closely connected among each other. It might be that network partitions join nodes with complementary characteristics, which are well adapted to varying environmental conditions (more like teams with complementary roles in different

environments). Group members that are specialized to specific environmental states might assure that their group achieves a high average fitness over time through these varying environments (some team members help the group thrive in some environments or time periods while others are champions in other environments and time periods). However, these are mere speculations and more in-depth research will be required to understand the reason for our quite pervasive finding.

The presented approach has several limitations that are worth noting. First, one limitation is that we only tested for two network partition algorithms (Fiedler's spectral and Kernighan-Lin), while we test for 30 different trait based types and up to 17 different network position based types per network. This makes our results with regard to network partitions less solid than our results about trait and network position metrics. A related issue that we do not understand yet is the limiting choice to work with rigid network partitions in our multilevel taxonomies. We chose this method to avoid confounding results in the descriptive analysis of the variance in fitness. However, it is very unlikely that natural and social hierarchies employ such mechanistic partitioning logic. It is much more likely that some kind of asymmetrical community structure is employed by evolutionary dynamics. Community structure algorithms (e.g. Girvan and Newman, 2002; Clauset et al. 2004; Bondel et al. 2008) detect more natural groups within a network than rigid partition algorithms. However, the result is groups of varying size, which makes them trickier for comparative exercises. We have not yet tested the effects of working with more flexible community detection algorithms to identify network structure.

Digging deeper it would also be interesting to see if different kinds of network metrics seem particularly apt at identifying large evolutionary forces and why. While we have shown that there is a statistically significant difference when thinking about evolutionary selection in terms of network metrics, the details of why which ones have what effect are far from being understood. For example, we find that closeness centrality metrics consistently produce notably extreme evolutionary forces. It might be that nodes that are on average close to all other nodes play a special role in evolutionary dynamics and are therefore subject to more selective pressure. More systematic work will need to be done to confirm and understand such indicative findings. This leads to an open research agenda on the analysis of traditional population dynamics from the perspective of network metrics with the aims of deepening our understanding of both evolutionary dynamics of social networks and the networked nature of evolution.

References

- Aldrich, H., & Ruef, M. (2006). *Organizations Evolving* (2nd ed.). Sage Publications Ltd.
- Bell, E. T. (1938). The Iterated Exponential Integrals. *Annals of Mathematics*, 39(3), 539–557.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *arXiv:0803.0476*. <http://doi.org/10.1088/1742-5468/2008/10/P10008>
- Boyd, R., & Richerson, P. J. (2005). *The Origin and Evolution of Cultures* (First Edition). Oxford University Press, USA.
- Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Harvard University Press. Retrieved from <http://papers.ssrn.com/abstract=1496205>
- Butts, C. T. (2008). A Relational Event Framework for Social Action. *Sociological Methodology*, 38(1), 155–200. <http://doi.org/10.1111/j.1467-9531.2008.00203.x>
- Cantor, M., Shoemaker, L. G., Cabral, R. B., Flores, C. O., Varga, M., & Whitehead, H. (2015). Multilevel animal societies can emerge from cultural transmission. *Nature Communications*, 6. <http://doi.org/10.1038/ncomms9091>
- Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Phys. Rev. E*, 70(66111). Retrieved from <http://arxiv.org/abs/cond-mat/0408187>
- Ewens, W. J. (1989). An interpretation and proof of the fundamental theorem of natural selection. *Theoretical Population Biology*, 36(2), 167–180. [http://doi.org/10.1016/0040-5809\(89\)90028-2](http://doi.org/10.1016/0040-5809(89)90028-2)
- Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2), 298–305.
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection* (1st ed.). Clarendon Press, Oxford.
- Frank, S. A. (1995). George Price's contributions to evolutionary genetics. *Journal of Theoretical Biology*, 175(3), 373–388. <http://doi.org/10.1006/jtbi.1995.0148>
- Frank, S. A. (1997). The Price Equation, Fisher's Fundamental Theorem, Kin Selection, and Causal Analysis. *Evolution*, 51(6), 1712–1729.
- Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton University Press. Retrieved from <http://stevefrank.org/foundations/foundations.html>
- Frank, S. A. (2012a). Natural selection. III. Selection versus transmission and the levels of selection. *Journal of Evolutionary Biology*, 25(2), 227–243. <http://doi.org/10.1111/j.1420-9101.2011.02431.x>
- Frank, S. A. (2012b). Natural selection. IV. The Price equation. *Journal of Evolutionary Biology*, 25(6), 1002–1019. <http://doi.org/10.1111/j.1420-9101.2012.02498.x>
- Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821–7826. <http://doi.org/10.1073/pnas.122653799>
- Haldane, J. B. S. (1924). A mathematical theory of natural and artificial selection—I. *Transactions of the Cambridge Philosophical Society*, 23, 19–24.
- Haldane, J. B. S. (1934). A Mathematical Theory of Natural and Artificial Selection Part X. Some Theorems on Artificial Selection. *Genetics*, 19(5), 412–429.

- Hannan, M. T., & Freeman, J. (1977). The Population Ecology of Organizations. *American Journal of Sociology*, 82(5), 929–964.
- Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Chung, S., Jimenez, J., ... Yildirim, M. A. (2011). *The atlas of economic complexity: mapping paths to prosperity*. Harvard University Center for International Development, MIT Media Lab. Retrieved from <http://atlas.media.mit.edu/>
- Hilbert, M. (2013). *Linking Information, Knowledge and Evolutionary Growth: A Multilevel Interplay between Natural Selection and Informed Intervention* (SSRN Scholarly Paper No. ID 2397751). Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=2397751>
- Hilbert, M. (2016). Formal definitions of information and knowledge and their role in growth through structural change. *Structural Change and Economic Dynamics*, <http://doi.org/10.1016/j.strueco.2016.03.004>.
- Jackson, M. O., & Watts, A. (2002). The Evolution of Social and Economic Networks. *Journal of Economic Theory*, 106(2), 265–295. <http://doi.org/10.1006/jeth.2001.2903>
- Jackson, M. O. (2010). *Social and Economic Networks*. Princeton University Press.
- Kernighan, B. W., & Lin, S. (1970). An Efficient Heuristic Procedure for Partitioning Graphs. *Bell System Technical Journal*, 49(2), 291–307. <http://doi.org/10.1002/j.1538-7305.1970.tb01770.x>
- Kossinets, G., & Watts, D. J. (2006). Empirical Analysis of an Evolving Social Network. *Science*, 311(5757), 88–90. <http://doi.org/10.1126/science.1116869>
- Lee, S., & Monge, P. (2011). The Coevolution of Multiplex Communication Networks in Organizational Communities. *Journal of Communication*, 61(4), 758–779. <http://doi.org/10.1111/j.1460-2466.2011.01566.x>
- Monge, P., Heiss, B. M., & Margolin, D. B. (2008). Communication Network Evolution in Organizational Communities. *Communication Theory*, 18(4), 449–477. <http://doi.org/10.1111/j.1468-2885.2008.00330.x>
- Nelson, R. R., & Winter, S. G. (1985). *An Evolutionary Theory of Economic Change*. Belknap Press of Harvard University Press.
- Newman, M. (2010). *Networks: An Introduction*. New York: Oxford University Press, USA.
- Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press of Harvard University Press.
- Palla, G., Barabási, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446(7136), 664–667. <http://doi.org/10.1038/nature05670>
- Price, G. R. (1970). Selection and Covariance. *Nature*, 227(5257), 520–521. <http://doi.org/10.1038/227520a0>
- Price, G. R. (1972a). Extension of covariance selection mathematics. *Annals of Human Genetics*, 35(4), 485–490. <http://doi.org/10.1111/j.1469-1809.1957.tb01874.x>
- Price, G. R. (1972b). Fisher's "fundamental theorem" made clear. *Annals of Human Genetics*, 36(2), 129–140.
- Snijders, T. A. B. (1996). Stochastic actor-oriented models for network change. *The Journal of Mathematical Sociology*, 21(1–2), 149–172. <http://doi.org/10.1080/0022250X.1996.9990178>

- Snijders, T. A. B. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology*, 31(1), 361–395. <http://doi.org/10.1111/0081-1750.00099>
- Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44–60. <http://doi.org/10.1016/j.socnet.2009.02.004>
- Stokman, F. N., & Doreian, P. (1997). Evolution of social networks: processes and principles. *Evolution of Social Networks*, 233–250.
- Stokman, F. N., & Doreian, P. (2001). Evolution of social networks Part II. Special Issue of *Journal of Mathematical Sociology*, 23(1).
- UNSD (United Nations Statistics Division). (2014). *National Classifications: 870 classifications in 154 countries*. New York. Retrieved from <http://unstats.un.org/unsd/cr/ctryreg/default.asp?Lg=1>

Supporting Information

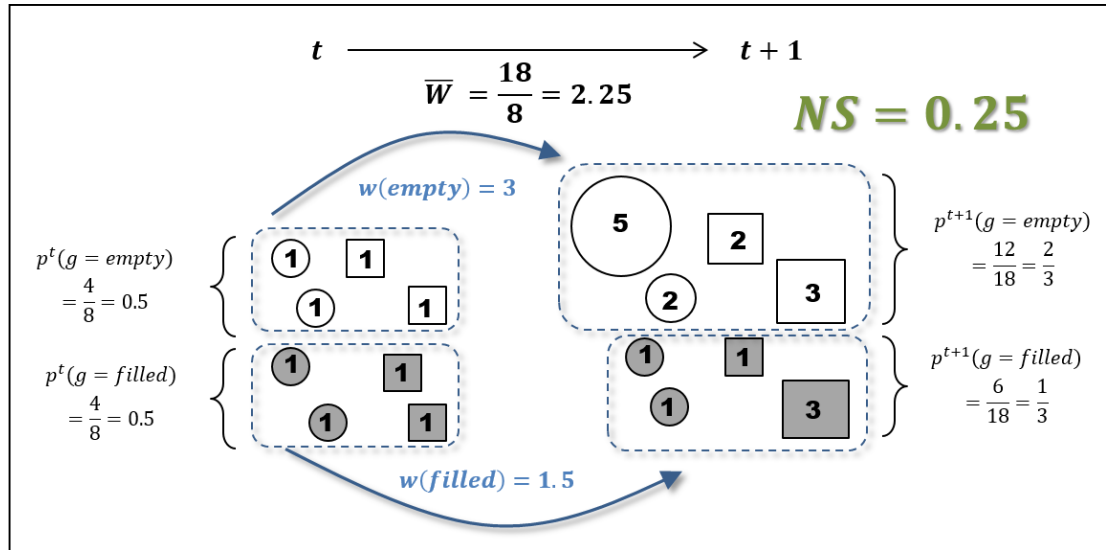
(A) Notation

We are aware that our notation is rather idiosyncratic, which is the result of deriving our notations from the equations of evolutionary population biology (i.e. Frank, 2012) and standard probability theory (to express multilevel hierarchies through nested conditioning). It is our hope that this compromise spreads the burden of adjustment among readers from different disciplines. At the same time, we are aware that this inevitably also results in the spread of discontent, for which we apologize in advance.

Figure S.1 is a more detailed version of Figure 1a from the main text aimed at illustrating our notation. The example tracks evolutionary updating between times t and $t + 1$. We have a random variable G (for “groups”) with two realizations (‘empty’ and ‘filled’), which has distribution $P^t(G)$ at time t and $P^{t+1}(G)$ at time t . For example, we have $p^t(g = \text{empty}) = 0.5$ and $p^{t+1}(g = \text{empty}) = \frac{2}{3}$. The letter w refers to the growth factor of fitness, $w = \left[\frac{\text{units}^{t+1}}{\text{units}^t} \right]$. The ‘empty’ type triples during updating, while the ‘filled’ type grows by 50 %: $w(g = \text{empty}) = 3$ and $w(g = \text{filled}) = 1.5$. The resulting population fitness \bar{W} is the expected value of the types’ fitness: $\bar{W} = \sum_g p(g) * w(g) = E_g[w] = 0.5 * 3 + 0.5 * 1.5 = 2.25$. We can now calculate the mean-normalized variance in fitness on this first level (natural selection NS).

$$\frac{Var(w)}{\bar{W}} = \frac{E[w]^2 - (E[w])^2}{\bar{W}} = \frac{(0.5 * 3^2 + 0.5 * 1.5^2) - (0.5 * 3 + 0.5 * 1.5)^2}{2.25} = 0.25 = NS$$

Figure S.1. Illustrative example of notation.



(B) Multilevel application

Following the multilevel logic outlined in equation (3) of the main article, we now calculate evolutionary forces within each of these groups on the second (more fine-grained) level. We redefine the random variable of the first level as $G_{l=1}$ (referring to ‘filling’) and introduce the additional random variable $G_{l=2}$ (referring to ‘color’). We then work with conditional probabilities (the lower level conditioned on the higher level in which it is nested). Figure S. 2 shows that the proportion of the red subtype on the second level $l = 2$ at time t represents one quarter of the empty types: $p^t(\text{red}_{l=2}|\text{empty}_{l=1}) = 0.25$ and that it multiplies by a factor of 5 during updating.

The type fitness on the higher level (aggregate coarse-grained) level is the expected value of the conditioned type fitness on the lower (more fine-grained) level. For the empty types: $w(\text{empty}_{l=1}) = \sum_{g_{l=2}} p^t(g_{l=2}|\text{empty}_{l=1}) * w(g_{l=2}|\text{empty}_{l=1}) = \frac{1}{4} * 5 + \frac{1}{4} * 2 + \frac{1}{4} * 2 + \frac{1}{4} * 3 = 3$.

We can now calculate the force of natural selection on the second level within the higher first level group:

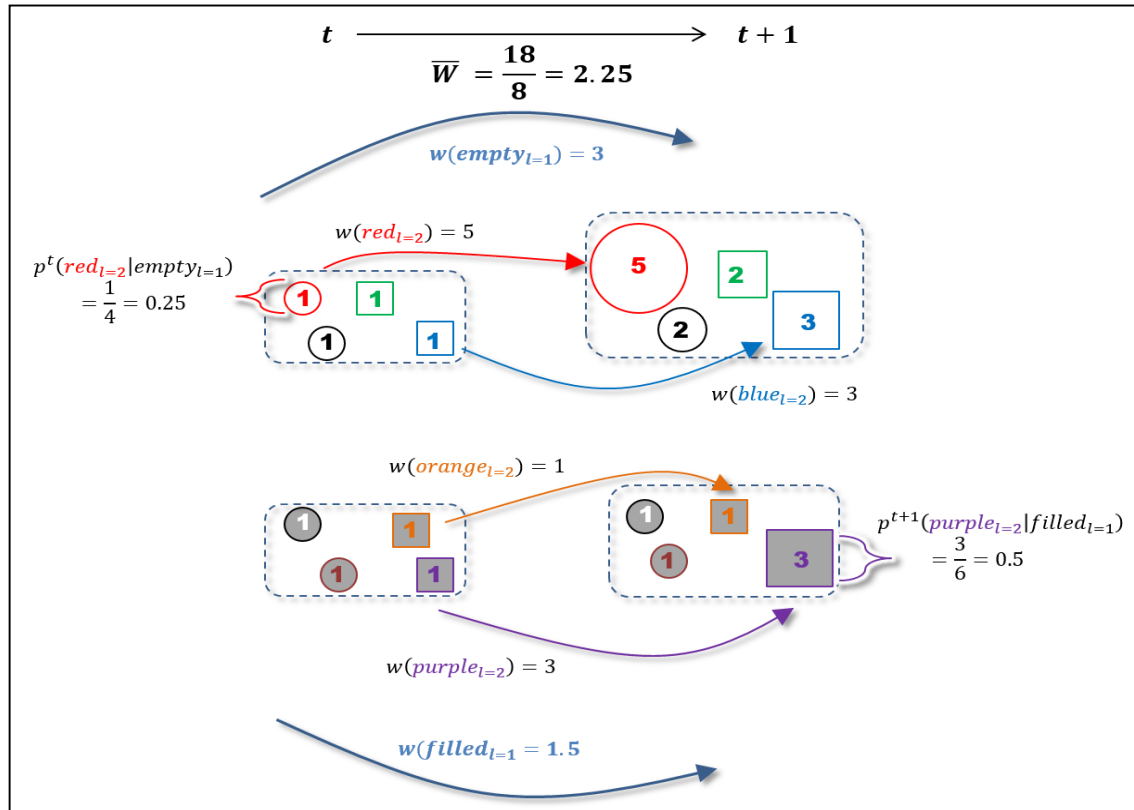
$$\frac{\text{Var}(w(g_{l=2}|g_{l=1}))}{w(g_{l=1})}$$

For the variance among colors within the empty types, this results in:

$$\begin{aligned} & \frac{\text{Var}(w(g_{l=2}|\text{empty}_{l=1}))}{w(\text{empty}_{l=1})} \\ &= \frac{\sum p(g_{l=2}|\text{empty}_{l=1}) * [w(g_{l=2}|\text{empty}_{l=1})]^2 - (\sum p(g_{l=2}|\text{empty}_{l=1}) * w(g_{l=2}|\text{empty}_{l=1}))^2}{w(\text{empty}_{l=1})} \\ &= \frac{(0.25 * 5^2 + 0.25 * 2^2 + 0.25 * 2^2 + 0.25 * 3^2) - (0.25 * 5 + 0.25 * 2 + 0.25 * 2 + 0.25 * 3)^2}{3} \\ &= 0.5. \end{aligned}$$

Coincidentally it turns out that likewise $\frac{\text{var}(w(g_{l=2}|\text{filled}_{l=1}))}{w(\text{filled}_{l=1})} = 0.5$. In this case, the second level $l = 2$ within each of the two groups on the higher level $l = 1$ is already the lowest level L permitted by the illustrative ‘database’ of our example. In our notation, this means $L = 2$. We can therefore calculate our metric of NS_{Total} (for details see Supporting Information C.3 below).

$$\begin{aligned} NS_{Total} &= \{NS_{l=1} + NS_{l=2}\} = \frac{\text{Var}(w(g_{l=1}))}{\bar{W}} + E_{l=1} \left[\frac{\text{Var}(w(g_{l=2}|g_{l=1}))}{w(g_{l=1})} \right] \\ &= 0.25 + [0.5 * 0.5 + 0.5 * 0.5] = 0.75 \end{aligned}$$

Figure S.2. Illustrative example of a second level decomposition of Figure S.1.

(C) Derivation of the multilevel decomposition

We use a three-part decomposition of evolutionary population change inspired by the Price equation (Price, 1970; 1972a; Frank, 1995; 1997; 2012). This kind of decomposition is useful in the quantitative comparison of different multilevel population partitions. The traditional Price equation (Price, 1970; 1972; Frank, 2012) keeps track of a chosen character (often represented by z , such as allele, genotype, phenotype, kind of individual or enterprise, etc.) and regresses this characteristic on the growth rate w . The selected character z has to be a scalar variable and is often subjectively defined by established scientific custom and/or current data availability. Since network groups do not have a quantitative scale and since we question the very nature of this choice of variable (*what variable defines groups in populations?*), we use the more fundamental version of Price's logic and set the character equal to fitness $z = w$. In this form, our decomposition is similar to the formulations used by evolutionary economists (Baily, et al., 1992; Foster, et al., 1998; Metcalfe and Ramlogan, 2006).

C.1 Three part decomposition: natural selection

We define delta Δ to be a difference by subtraction in a time series: $\Delta w = w^{t+1} - w^t$ and $\Delta p = p^{t+1} - p^t$. We decompose the change in population fitness $\Delta \bar{W}$ (on the highest level). In the following reformulation, we make use of the fact that $p(g)^{t+1} = p(g)^t * \frac{w(g)}{\bar{W}}$ (which is known as the 'replicator equation' in economics and game theory). For reasons of simplicity, the superscript t is omitted in the

following derivation when the variable refers to the initial generation, e.g. $p = p^t$. Likewise, in one-level decompositions the reference to the random variable is omitted, $p = p(g)$.

$$\begin{aligned}
 \Delta \bar{W} &= \bar{w}^{t+1} - \bar{w}^t = \sum_g p(g)^{t+1} * w(g)^{t+1} - \sum_g p(g) * w(g) \\
 &= \sum_g (p + \Delta p) * (w + \Delta w) - \sum_g p * w \\
 &= \sum_g [(p * w) + (\Delta p * w) + (p * \Delta w) + (\Delta p * \Delta w)] - \sum_g p * w = \\
 &= \sum \Delta p * w + \sum p * \Delta w + \sum \Delta p * \Delta w \tag{S.1} \\
 &= \sum (p * \frac{w}{\bar{W}} - p) * w + \sum p * \Delta w + \sum (p^{t+1} - p) * \Delta w = \\
 &= \left[\sum p * \frac{(w)^2}{\bar{W}} - \sum (p * w) * \frac{\bar{W}}{\bar{W}} \right] + \left[\sum p * \Delta w \right] + \left[\sum p^{t+1} * \Delta w - 1 * \sum p * \Delta w \right] = \\
 &= \frac{1}{\bar{W}} \left[\sum p * w^2 - \left(\sum p * w \right)^2 \right] + \left[\sum p * \Delta w \right] + \left[\sum p \frac{w}{\bar{W}} * \Delta w - \sum p^{t+1} * \sum p * \Delta w \right] = \\
 &= \frac{1}{\bar{W}} Var(w) + \left[\sum p \frac{w}{\bar{W}} * \Delta w - \sum p \frac{w}{\bar{W}} * \sum p \Delta w \right] + \left[\sum p * \Delta w \right] \\
 \Delta \bar{W} &= \frac{Var(w)}{\bar{W}} + Cov\left(\frac{w}{\bar{W}}, \Delta w\right) + \sum p * \Delta w = NS + RA + MLT \tag{S.2}
 \end{aligned}$$

While the simplicity of equation (S.1) makes it most useful for empirical calculations, the equivalent reformulation of equation (S.2) facilitates meaningful interpretation of the three parts. The first component, $\frac{Var(w)}{\bar{W}}$, refers to natural selection (*NS*), in line with Fisher's fundamental theorem of natural selection (Fisher, 1930; Price, 1972b; Ewens, 1989; Lessard, 1997; Frank, 1997; Edwards, 2002). The theorem states that the change in average fitness *caused by natural selection* is proportional to the variance in fitness. Since both, the variance and its normalizing population fitness can only be positive, the corresponding contribution to change in population fitness (left hand side) can only be positive:

$$\frac{Var(w)}{\bar{W}} \geq 0 \tag{S.3}; \text{ Fisher's fundamental theorem of natural selection}$$

Since Fisher's fundamental theorem only refers to the part of the total change that is caused by natural selection, it has been referred to as a "partial change" in fitness by biologists (Frank, 1998). Economists have referred to this incomplete modeling of evolutionary dynamics as a restricted model of "pure selection" (Nelson and Winter; 1985; p. 243).

C.2 Two part decomposition: sustainable selection

Adding up the variance and the covariance terms, we can derive equation (2) from the main article:

$$\begin{aligned}
\Delta \bar{W} &= \frac{Var(w)}{\bar{w}} + Cov\left(\frac{w}{\bar{w}}, \Delta w\right) + \left\{ \sum p * \Delta w \right\} = \text{accrd. to (S.1)} \\
&= \sum \Delta p * w + \left[\sum \Delta p * \Delta w \right] + \left\{ \sum p * \Delta w \right\} = \\
&= \sum \Delta p * w + \left[\sum \Delta p * w^{t+1} - \sum \Delta p * w \right] + \left\{ \sum p * \Delta w \right\} \quad (S.4.1) \\
&= \sum p^{t+1} * w^{t+1} - \sum p * w^{t+1} + \left\{ \sum p * \Delta w \right\} = \sum p \frac{w}{\bar{w}} * w^{t+1} - 1 * \sum p * w^{t+1} + \left\{ \sum p * \Delta w \right\} \\
&= \sum p \frac{w}{\bar{w}} w^{t+1} - \sum p \frac{w}{\bar{w}} * \sum p * w^{t+1} + \left\{ \sum p * \Delta w \right\} \\
&= Cov\left(\frac{w}{\bar{w}}, w^{t+1}\right) + \left\{ \sum p * \Delta w \right\} = \frac{Cov(w^t, w^{t+1})}{\bar{w}} + \left\{ \sum p * \Delta w \right\} \\
&= SS + MLT \quad (S.4) \text{ Sustainable Selection}
\end{aligned}$$

As shown by the canceling terms in equation (S.4.1), the Relative Acceleration (*RA*) term incorporates a negative counterpart of the positive contributions of Natural Selection (*NS*) to increases in fitness. This is why we see the “mirror result” between these two terms when applied to empirical data (Hilbert, 2013), with the Relative Acceleration (*RA*) being a negative counterpart to Natural Selection (*NS*).

C.3 Multilevel decomposition: nesting as conditional proportions

We now decompose the multilevel-term (*MLT*) of equation (S.2) through recursion with the same logic, just one level lower. In order to be able to work with an arbitrary number of level, we introduce a scalable notation for levels which is indexes with the subscript l . Population fitness on the highest possible level (which is traditionally expressed with an overbar \bar{W}) is found on level $l = 0$ in terms of this scalable multilevel notation: $\bar{W} = w(g_{l=0})$. The types of the first level are indexed with the subscript $l = 1$. Therefore, $p(g_{l=1})$ represents a proportion on a first level group within the population (more precise: $p(g_{l=1}|g_{l=0})$, proportion of type, conditioned on the entire population). We define $P(g_{l=2}|g_{l=1})$ as the conditional distribution of a subpopulation on level $l = 2$ conditioned on the higher level group at $l = 1$, which has the same properties as conditional probabilities, with $\sum_{g_{l=2}} p(g_{l=2}|g_{l=1}) = 1$. In other words, the index $(g_{l=2}|g_{l=1})$ tracks the number of subtypes on level $l = 2$ within a specific higher level type on level on level $l = 1$ (which is implicitly understood to be part of the total population on level $l = 0$). Expressing equation (S.2) in this more detailed notation we get:

$$\Delta W(g_{l=0}) = \frac{Var(w(g_{l=1}))}{w(g_{l=0})} + \sum_{g_{l=1}} p(g_{l=1}) * [\Delta w(g_{l=1})] + Cov\left(\frac{w(g_{l=1})}{w(g_{l=0})}, \Delta w(g_{l=1})\right) =$$

$$\begin{aligned}
&= \frac{Var(w(g_{l=1}))}{w(g_{l=0})} \\
&\quad + \sum_{g_{l=1}} p(g_{l=1}) \left[\frac{Var(w(g_{l=2}|g_{l=1}))}{w(g_{l=1})} + \sum_{g_{l=2}} p(g_{l=2}|g_{l=1}) * [\Delta w(g_{l=2}|g_{l=1})] \right. \\
&\quad \left. + Cov\left(\frac{w(g_{l=2}|g_{l=1})}{w(g_{l=1})}, \Delta w(g_{l=2}|g_{l=1})\right) \right] + Cov\left(\frac{w(g_{l=1})}{w(g_{l=0})}, \Delta w(g_{l=1})\right) = \\
&= \left\{ \frac{Var(w(g_{l=1}))}{w(g_{l=0})} + \sum_{g_{l=1}} p(g_{l=1}|g_{l=0}) * \frac{Var(w(g_{l=2}|g_{l=1}))}{w(g_{l=1})} \right\} \\
&\quad + \left\{ \sum_{g_{l=1}} p(g_{l=1}|g_{l=0}) \sum_{g_{l=2}} p(g_{l=2}|g_{l=1}) * [\Delta w(g_{l=2}|g_{l=1})] \right\} \\
&\quad + \left\{ Cov\left(\frac{w(g_{l=1})}{w(g_{l=0})}, \Delta w(g_{l=1})\right) + \sum_{g_{l=1}} p(g_{l=1}|g_{l=0}) * Cov\left(\frac{w(g_{l=2}|g_{l=1})}{w(g_{l=1})}, \Delta w(g_{l=2}|g_{l=1})\right) \right\} \\
&= \left\{ \frac{Var(w(g_{l=1}))}{w(g_{l=0})} + E_{l=1} \left[\frac{Var(w(g_{l=2}))}{w(g_{l=1})} \right] \right\} \\
&\quad + \left\{ Cov\left(\frac{w(g_{l=1})}{w(g_{l=0})}, \Delta w(g_{l=1})\right) + E_{l=1} \left[Cov\left(\frac{w(g_{l=2})}{w(g_{l=1})}, \Delta w(g_{l=2})\right) \right] \right\} \\
&\quad + \left\{ \sum_{g_{l=2}} p(g_{l=2}) * \Delta w(g_{l=2}) \right\} \quad (S.5)
\end{aligned}$$

$$\Delta W(g_{l=0}) = \{NS_{l=1} + NS_{l=2}\} + \{RA_{l=1} + RA_{l=2}\} + \{MLT_{l=2}\}$$

This logic can be continued over further levels of fine-graining, indexed by l , with $l = [1, 2, 3, \dots, L]$, until the lowest level L permitted by the available data.⁴ This makes the conditional logic scalable. At the end we are left with three terms (the three terms in the {braces}). The terms within the braces can be summed up from the first to the lowest level, always leaving the multilevel term on the lowest level permitted by the database. This is represented by:

$$\Delta w(g_{l=0}) = NS_{Total} + RA_{Total} + MLT_{l=L}$$

For example, in case there are only two levels (such as in the example of Figure S.2 from above), the content of the three terms inside {braces} in equation (S.5) equation provide the details of this representation in totals. We can expand this logic to an arbitrary number of levels until level L . The multilevel total of each evolutionary force essentially results in a sum of (variance or covariance) terms over the consecutive levels, while these terms are weighted with the product of conditional proportions over the respective levels (which, according to the chain rule, essentially results 'joint probabilities' on the corresponding level):

⁴ With $L \leq \log_2(\text{number of smallest units in database})$, since the smallest possible group consists of two members.

$$\Delta W(g_{l=0}) = \sum_{l=0}^{l=L} \prod_{l=0}^l p(g_{l+1}|g_l) * \frac{Var(w(g_{l+1}))}{w(g_l)} + \quad (S.6.1; NS_{Total})$$

$$+ \sum_{l=0}^{l=L} \prod_{l=0}^l p(g_{l+1}|g_l) * Cov\left(\frac{w(g_{l+1})}{w(g_l)}, \Delta w(g_{l+1})\right) + \quad (S.6.2; RA_{Total})$$

$$+ \sum_{g_{l=L}} p(g_{l=L}) * \Delta w(g_{l=L}) \quad (S.6.3; MLT_{l=L})$$

The same holds for the SS_{Total} term.

$$\begin{aligned} \Delta W(g_{l=0}) &= SS_{Total} + MLT_{l=L} \\ &= \sum_{l=0}^{l=L} \prod_{l=0}^l p(g_{l+1}|g_l) * \frac{Cov(w(g_{l+1})^t, w(g_{l+1})^{t+1})}{w(g_l)} + E_L[\Delta w(g_L)] \quad (S.7) \end{aligned}$$

(D) Proof of applicability of comparative method

Here we show the proof of equation (4) from the main article, namely that the sum of the variance (NS_{Total}) and covariance terms (RA_{Total}) add up to the same combined sustainable selection (SS_{Total}) effect for all different possible ways to structure group membership, given the condition that this multilevel logic is pursued until the lowest possible level of fine-graining L . The proof follows directly from equation (S.7) from above. Saying that $[NS_{Total} + RA_{Total} = SS_{Total} = \text{constant}]$ for all possible partitions is equivalent to saying that the remaining multilevel term on the lowest level MLT_L is constant for all possible partitions. This is always the case, because partitions on the lowest level have equal proportions. Shares on the lowest level, $p(g_{l=L})$, are defined by the permitted level of fine-graining of the database, which is independent from different ways of higher level grouping of this lowest level granularity. For example, in our example of Figure S.2, the lowest (second level) distribution of the updated period at $t + 1$ is $P^{t+1}(g_{l=2=L})$ consists of the proportions $\left\{\frac{5}{18}; \frac{2}{18}; \frac{2}{18}; \frac{3}{18}; \frac{1}{18}; \frac{1}{18}; \frac{1}{18}; \frac{3}{18}\right\}$ (corresponding to the colored second level subtypes from Figure S.2), which is independent from different ways to create first level types, such as done in Figure S.1. In formal terms of our notation:

$$\begin{aligned} MLT_L &= \sum_L \prod_{l=0}^L p(g_{l+1}|g_l) * [\Delta w(g_L)] = \sum_L p(g_L) * [\Delta w(g_L)] = E_L[\Delta w(g_L)] \\ &= \text{constant for all partitions} \end{aligned}$$

(E) Summary of datasets

The following Table list the different attribute and network metrics used to partition populations.

	Democracy Now	Nova PBS	Kiva Leader Growth variable: - % of total raised - abs. USD raised	Kiva Team Growth variable: - % of total raised - abs. USD raised	Intern. Trade Growth variable: - national total - per capita
# Nodes	56	86	120	118	118
# Attributes	10	10	9	10	30
# Netw.metr.	10	11	16	17	12 (nat.); 14 (per capita)
Classification Attribute	Likes	Likes	NoBorrowers	#ofBorrowers	Population
	LikesPerView	Dislikes	PartnerDefaultRate	RepaymentTerm	Total natural resources rents (%GDP)
	Dislikes	Likes-Dislikes	PartnerDelinquencyRate	PartnerTotLoans	School enrollment, tertiary (% gross)
	DislikesPerView	Data Published	Partner#Entrepre	PartnerTimeOnKiva	School enrollment, primary (% gross)
	Likes-Dislikes	Comment	PartnerLoansAtRisk	Partner#Entrepr	Scientific&technical journal articles
	Likes-DislikesPerView	CommentPERview	PartnerTimeOnKiva	PartnerAvgLoanSize	Govmn.consumpt. exp. (const.'05\$)
	DatePublished	LikesPERview	PartnerTotalLoans	PartnerDelinquRate	Govmn.consumpt. exp. (%GDP)
	Comments	DislikesPERview	RepaymentTermMonths	PartnerLoansAtRisk	Export-Import
	CommentsPerView	Likes-DislikesPERview	GrowthOfGrowthVar	PartnerDefaultRate	Extrn.balance goods&serv.(%GDP)
	GrowthOfGrowthVar	GrowthOfGrowthVar		GrowthOfGrowthVar	Avg. precipitation in depth (mm/year)
					Arable land (hectares)
					Arable land (hectares/person)
					consumpt. exp. (const.'05\$)
					consumpt. exp. (%GDP)
					Exports goods&serv. (const.'05\$)
					Exports goods&serv. (%GDP)
					Imports goods&serv. (const.'05\$)
					Imports goods&serv. (%GDP)
					Fertility rate (births/woman)
					Gross capital formation (const.'05\$)
					Gross capital formation (%GDP)
Classification Network Metric	EigenvectorCentrality	Centr.Eigenvector	Capability	Capability	Centr.Authority
	Centr.Eigenv.PerComp.	Centr.Eigenv.PerComp.	Centr.Authority	Centr.Authority	Centr.Closeness/inv.=1/[unsc.]
	CentralityAuthority	Centr.Authority	Centr.Bonacich Power	Centr.Betweenness	Centr.Eigenvector
	CognitiveDemand	Centr.In-Closeness	Centr.Closeness	Centr.Bonacich Power	Centr.Eigenv.PerComp.
	CentralityTotalDegree	Cognitive Demand	Centr.Eigenvector	Centr.Closeness	Centr.Hub
	Centr.Column Degree	Centr.Total Degree	Centr.Information	Centr.Eigenvector	Centr.In Degree/[unsc.]
	ColumnGiniMeansDiff.	Centr.Column Degree	Centr.Total Degree	Centr.Hub	Centr.In-Closeness/inv.=1
	CentralityInDegree	Column Gini Means Diff.	Clique Count	Centr.Total Degree	Centr.Information
	DensityClusteringCoef.	Centr.In Degree	Cognitive Demand	Clique Count	Centr.Information/[unsc.]
	CliqueCount	Clique Count	Column Gini Means Diff.	Cognitive Demand	Centr.Inverse Closeness/inv.=1
		Density Clustering Coef.	Constraint Burt	Column Gini Means Diff.	Centr.Out Degree/[unsc.]
			Correlation Similarity	Constraint Burt	Centr.Total Degree/[unsc.]
			Density Clustering Coef.	Correlation Similarity	Centr.In Degree (only PerCapita)
			Effective Netw. Size Burt	Density Clustering Coef.	Centr.Inv. Closeness (only PerCapita)
			Simmelian Ties	Effective Netw. Size Burt	
			Triad Count	Simmelian Ties	
				Triad Count	
Network Partition	Kernighan-Lin partition	Kernighan-Lin partition	Kernighan-Lin partition	Kernighan-Lin partition	Kernighan-Lin partition
	Fiedler's Spectral partition	Fiedler's Spectral partition	Fiedler's Spectral partition	Fiedler's Spectral partition	Fiedler's Spectral partition

DemocracyNow: YouTube video views of news videos

A webscraper was programmed to obtain this dataset from the YouTube.com video news outlet of the nonprofit and independently syndicated news program “Democracy Now!” (hosted by Amy Goodman and Juan Gonzalez): <http://www.youtube.com/user/democracynow/videos?sort=p&flow=list&view=0>. The program is funded entirely through contributions from listeners, viewers, and foundations, and does not accept advertisers, corporate underwriting, or government funding. The 29 “most popular” (most viewed) videos were selected on July 24, 2013 and tracked every 12 hours for the consecutive 12 days at 5am and 5pm, until August 5, 2013, resulting in 24 time steps. Clicking on each of these 29 videos, the YouTube platform provides links to some 20 additional videos at the right-hand side of the screen. We collected several attributes of the 29 videos + the linked videos. The nodes of the resulting network are the videos and directed links are established by the hyperlinks to other videos provided at the right-hand side of the video display. We eliminated those videos (nodes) that did not appear at each of the 24 periods. The result was a network with 165 nodes. We summed up the resulting 24 networks to create one union network (a tie-strength of 24 would mean that those two videos are linked at each of the 24 periods⁵). After testing for several cut-off points, we eliminated those links that did not achieve a tie-strength of 15 and above (the number of nodes in the resulting network stayed the same until a tie-strength of 21 = 87.5 % of the periods, which justifies this cut-off). The resulting network of 148 nodes contains many pendants (nodes only connected to one other node), which leads to confounding results in our evolutionary dynamics, since those video views are influenced by references to videos outside of our sample. After eliminating the pendants we are left with one directed union network of 56 nodes, and reintroduced the full tie-strength between 1 and 24 for those nodes. This includes the 29 most popular videos on the channel (our primary source) + 27 videos that were regularly linked to those popular videos (these videos might or might not be from the DemocracyNow channel and include videos from other news channels, individuals and bloggers, and entertainment provides, among others). The main evolutionary growth variable consists of the number of views of the video. We also collected the number of Likes and Dislikes, the date published and the number of comments. We sum these indicators up for the 24 periods and ranked them from lowest to highest to create one attribute scale for the attribute partitioning.

Figure S-1: DemocracyNow union network (56 nodes)



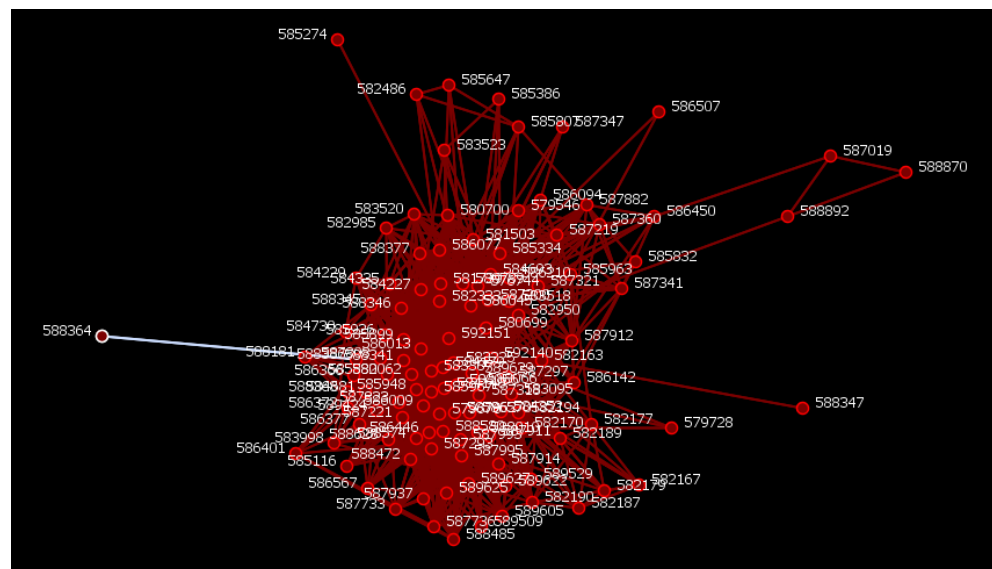
Source: own elaboration, display with ORA software, CASOS, Carnegie Mellon University.

⁵ Tie-strengths above 24 sometimes occur and indicate that the same video was posted more than once for one time-instance as a link by the YouTube algorithm.

KIVA individual lenders: crowd-sourced entrepreneur fundraising

A webscraper was programmed to obtain this dataset from Kiva Microfunds (Kiva.org), a non-profit organization that allows people to lend money via the Internet to underserved entrepreneurs and students in 70 countries: <http://www.kiva.org/lend?sortBy=amountLeft>. On August 6, 2013 the 189 entrepreneurs with the “largest amount left” (in % of total amount = minimum raised) were selected and the evolution of their fundraising efforts were recorded every 6 hours for the consecutive 6 days at 1am, 7am, 1pm and 7pm, until August 12, 2013, resulting in 24 time steps. The nodes in the networks are the entrepreneurs and we establish an undirected link between two entrepreneurs if the same borrower has lent funds to both entrepreneurs. Each shared lender adds an additional tie-strength of 1. We summed up the resulting 24 networks to create one union network. We kept the nodes that made part of the resulting giant component (largest connected subgraph) and removed pendants, leaving us with a connected subgraph of 120 nodes for the “Kiva lender network”. We realize the exercise for two different growth variables: (A) the amount raised as percentage of total fundraising goal; (B) the amount raised in total US\$. We collect different attributes of the (teams of) entrepreneurs (for example on the number of borrowers that teamed up for one project and on the characteristics of the field partner of the entrepreneurs, which can be microfinance institutions, social businesses, schools or non-profit organizations). We sum these indicators up for the 24 periods and rank them from lowest to highest to create one attribute scale for the attribute partitioning.

Figure S-3: Kiva lender network (120 nodes)



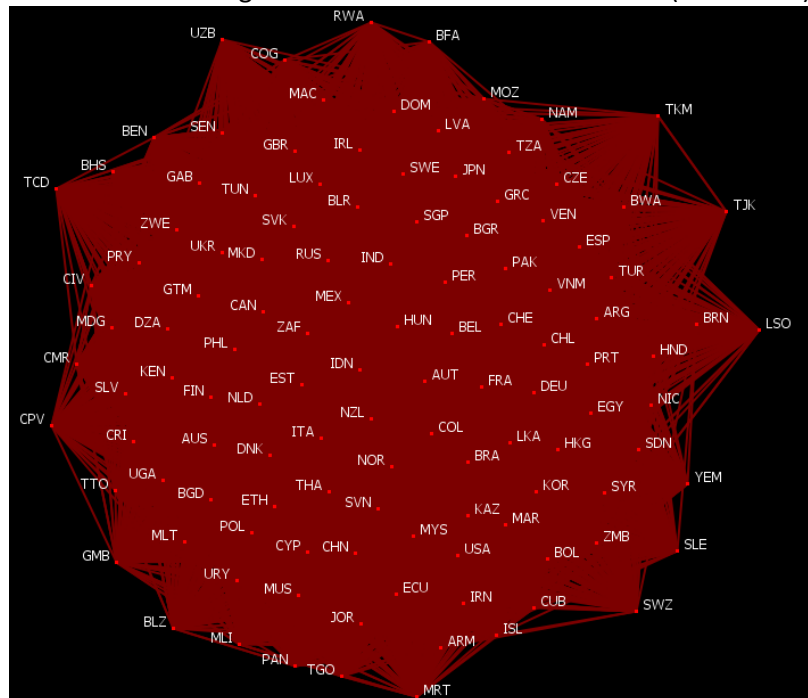
Source: own elaboration, display with ORA software, CASOS, Carnegie Mellon University.

KIVA teams: crowd-sourced entrepreneur fundraising

Lenders on Kiva.org can also join a team. There are local areas teams (with members from the same country or region), business teams (with members from the same company), religious congregations, sports clubs, or simply friends, etc. We also collected those and created a second Kiva network. In this network the nodes have the same attributes as in the “Kiva lender network”, but the network ties are teams, and not individual donors. After removing the pendants we are left with 118 nodes for the “Kiva

team network". We summed up the resulting 24 networks to create one union network. The attribute measures are the same as in the KIVA individual lenders network.

The evolutionary growth variable of this dataset consists of Gross National Product (GDP) in constant US \$ (of 2005) and for another test in GDP per capita of importing country in constant US \$ (of 2005) (World Bank, 2011). We created a network of the sum of all international trade flows for each year between 1987 and 2010 (in current US dollars) (based on UN Comtrade SITIC4rev2, which was cleaned and kindly shared with us by Hausmann, Hidalgo, Bustos, et al., 2011). We summed up the resulting 24 networks to create one union network. As classifying attribute measures we used several World Development indicators (World Bank, 2011). We sum them up for the 24 periods and rank them from lowest to highest to create one attribute scale for the attribute partitioning. Given the limitation in available time series for World Development Indicators, we had to reduce the network to 118 country nodes, which represent 97 % of global GDP and 91 % of the global population.

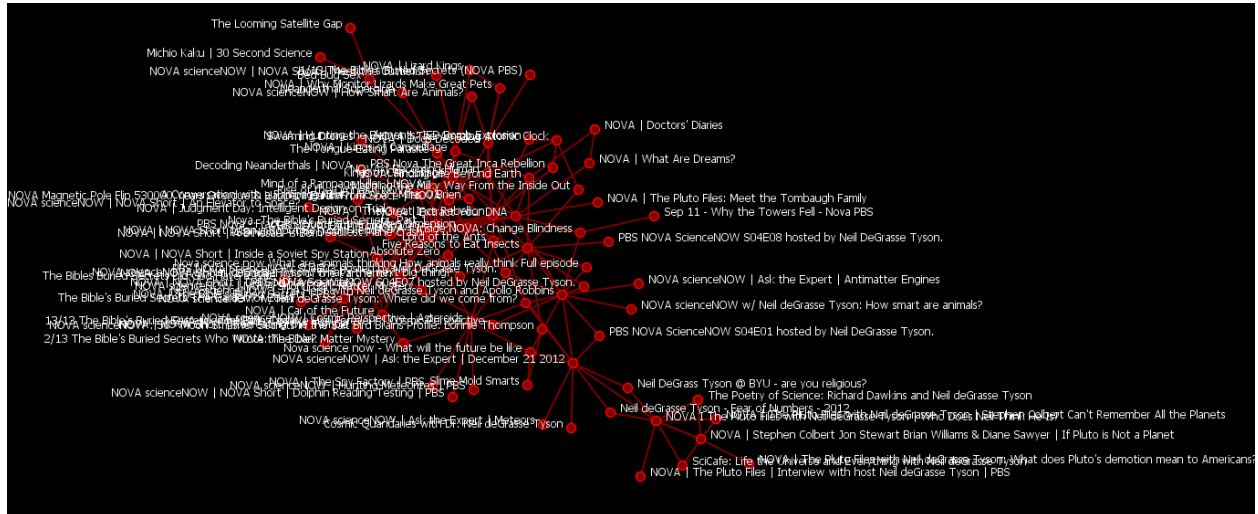


Source: own elaboration, display with ORA software, CASOS, Carnegie Mellon University.

The dataset preparation follows the same procedure as for the above described DemocracyNow dataset, being scraped from the YouTube.com video outlet of the program NOVA of U.S.'s Public Broadcasting Service (PBS), during the same period, but collected at 6am and 6pm: <http://www.youtube.com/user/NOVAonline/videos?flow=list&view=0&sort=p>. Since the link structure was more volatile in this case, we only counted links with a tie-strength of and above 22 (out of 24) in the

summed union network.⁵ Before the elimination of pendants, the result was a union network of 156 nodes, and we worked with the directed network of 86 nodes that was the result after the elimination of pendants. We then reintroduced the full tie-strength of 1-24 for the identified 86 nodes.

Figure S-2: NOVA PBS union network (86 nodes)



Source: own elaboration, display with ORA software, CASOS, Carnegie Mellon University.

References

- Baily, M. N., Hulten, C., Campbell, D., Bresnahan, T., & Caves, R. (1992). Productivity Dynamics in Manufacturing Plants. *Brookings Papers on Economic Activity. Microeconomics*, 1992, 187–267.
- Calcul. (2012). *Bell Number Approximation Calculator*. Retrieved from <http://www.calcul.com/bell-number-approximation>
- Edwards, A. W. F. (2002). The Fundamental Theorem of Natural Selection. *Theoretical Population Biology*, 61(3), 335–337. <http://doi.org/10.1006/tpbi.2002.1570>
- Ewens, W. J. (1989). An interpretation and proof of the fundamental theorem of natural selection. *Theoretical Population Biology*, 36(2), 167–180. [http://doi.org/10.1016/0040-5809\(89\)90028-2](http://doi.org/10.1016/0040-5809(89)90028-2)
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection* (1st ed.). Clarendon Press, Oxford.
- Foster, L., Haltiwanger, J., & Krizan, C. J. (1998). Aggregate Productivity Growth: Lessons from Microeconomic Evidence. *National Bureau of Economic Research Working Paper Series, No. 6803*. Retrieved from <http://www.nber.org/papers/w6803>
- Frank, S. A. (1995). George Price's contributions to evolutionary genetics. *Journal of Theoretical Biology*, 175(3), 373–388. <http://doi.org/10.1006/jtbi.1995.0148>
- Frank, S. A. (1997). The Price Equation, Fisher's Fundamental Theorem, Kin Selection, and Causal Analysis. *Evolution*, 51(6), 1712–1729.
- Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton University Press. Retrieved from <http://stevefrank.org/foundations/foundations.html>

- Frank, S. A. (2012). Natural selection. IV. The Price equation. *Journal of Evolutionary Biology*, 25(6), 1002–1019. <http://doi.org/10.1111/j.1420-9101.2012.02498.x>
- Frank, S. A. (2013). Natural selection. VI. Partitioning the information in fitness and characters by path analysis. *Journal of Evolutionary Biology*, 26(3), 457–471. <http://doi.org/10.1111/jeb.12066>
- Hilbert, M. (forthcoming). Linking Information, Knowledge and Evolutionary Growth: A multilevel interplay between natural selection and informed intervention. *Preprint Version, Presented at “Advances in Dynamic Economics and Development: Economic and Complexity”, Nov. 2013*. Retrieved from <http://www.economia.ufpr.br/Eventos/Complexity2013/Hilbert%20-%20Linking%20information%20, knowlege%20and%20evolutionary%20Growth.pdf>
- Lessard, S. (1997). Fisher’s Fundamental Theorem of Natural Selection Revisited. *Theoretical Population Biology*, 52(2), 119–136. <http://doi.org/10.1006/tpbi.1997.1324>
- Metcalfe, J. S., & Ramlogan, R. (2006). Creative Destruction and the Measurement of Productivity Change. *Revue de l’OFCE*, 97 bis(5), 373–397.
- Nelson, R. R., & Winter, S. G. (1985). *An Evolutionary Theory of Economic Change*. Belknap Press of Harvard University Press.
- Price, G. R. (1970). Selection and Covariance. *Nature*, 227(5257), 520–521. <http://doi.org/10.1038/227520a0>
- Price, G. R. (1972a). Extension of covariance selection mathematics. *Annals of Human Genetics*, 35(4), 485–490. <http://doi.org/10.1111/j.1469-1809.1957.tb01874.x>
- Price, G. R. (1972b). Fisher’s “fundamental theorem” made clear. *Annals of Human Genetics*, 36(2), 129–140.
- Schumpeter, J. (1939). *Business Cycles: A Theoretical, Historical, And Statistical Analysis of the Capitalist Process*. New York: McGraw-Hill. Retrieved from http://classiques.uqac.ca/classiques/Schumpeter_joseph/business_cycles/schumpeter_business_cycles.pdf
- Simon, H. A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482.
- Wikipedia. (2013). Bell number. Retrieved from http://en.wikipedia.org/wiki/Bell_number