doi: 10.1038/nature06940

Social diversity promotes the emergence of cooperation in public goods games

Francisco C. Santos[1], Marta D. Santos[2] & Jorge M. Pacheco[2]

[1] *IRIDIA*, Computer and Decision Engineering Department,

Université Libre de Bruxelles, Brussels, Belgium,

[2] *ATP*-group, CFTC, and Departamento de Fisica da Universidade de Lisboa,

Complexo Interdisciplinar, Av Prof. Gama Pinto 2, 1649-003 Lisboa, Portugal

SUPPLEMENTARY INFORMATION

15 pages and 8 figures.

Evolutionary dynamics of the Rich, the Poor, the Marginally and the Centrally connected

When population structure is associated with a heterogeneous social graph, one introduces diversity in the roles each individual plays in the community. In particular, when populations are structured following a scale-free distribution, the majority of individuals engage only in a few games, while a minority is able to participate in many games[27]. Depending on the size and number of the PGG each individual participates, the payoffs of an individual can rise to high values with respect to the population average, or may remain very low. In other words, the diversity introduced leads to another type of diversity which can be deduced from the fitness (or wealth, in a more economical sense) distribution. It is noteworthy, however, that the nature of the public good games makes the income of an individual depend not only on her number of social ties (her degree in the social graph), but also on the "degree" of her neighbors. In this sense, heterogeneous graphs lead to the appearance of several classes of individuals, both in what concerns the number of games in which they participate and also in what concerns their wealth.

Let us consider for simplicity, three classes of individuals regarding their number of social ties k_i on a scale-free network:

i) Low-Degree class, whenever $k_i < z$;

ii) Medium-Degree class, whenever $z \le k_i < \frac{k_{\max}}{3}$ and finally

iii) High-Degree class, whenever $\frac{k_{\max}}{3} \le k_i \le k_{\max}$,

encompassing the highly connected individuals (the hubs).

Let us also distinguish individuals in another group of three classes based on their personal wealth (fitness) Π_i:

i) Lower-Class, when $\Pi_i < \frac{\Pi_{total}}{3}$;

ii) Middle-Class, when $\frac{1}{3}\Pi_{total} \le \Pi_i < \frac{2}{3}\Pi_{total}$ *and*

iii) Upper-Class, when $\frac{2}{3}\Pi_{total} \le \Pi_i \le \Pi_{total}$,

where Π_{total} is the total wealth of the population at a given time of the evolutionary process.

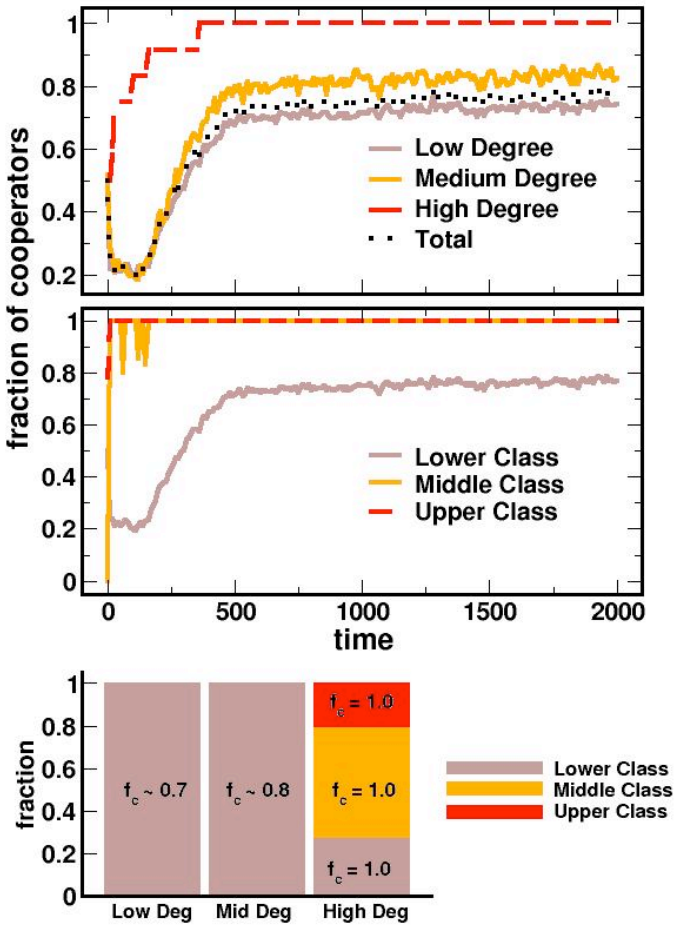

Figure S1. **Time-dependence of the fraction of cooperators on scale-free communities.**

In Figure S1 we plot the fraction of cooperators in the population for each of these 9 classes during evolution, when each individual invests the same amount irrespective of the number of PGG she participates. The time-dependent curves in Figure S1

provide a representative *run* for a multiplication factor $r=1.7$, that is, $\eta=0.34$ ($z=4$). For this value of η, cooperators dominate the population while unable to wipe out defectors.

In the upper panel, it is clear that individuals of High-Degree quickly become cooperators and remain so for the rest of the evolutionary process. On the contrary, Low-Degree and Medium-Degree classes start by adopting defector strategies. The more Low-Degree and Medium-Degree individuals adopt the **D** strategy, the more vulnerable they become to the influential role played by **C** hubs (High-Degree individuals). Consequently, the situation quickly reverts to a scenario in which **D**s survive only as individuals of Low-Degree and Medium-Degree.

The middle panel shows that high levels of wealth (fitness) – Upper-Class – are associated with **C**s. Moreover, no **D**s survive in the Middle-Class, being all relegated to the Lower-Class.

Finally, the lower panel summarizes together the information provided separately in the upper panels, correlating fitness, degree and strategy. Only individuals of High-Degree classes can achieve the Upper-Class in terms of wealth. On the other hand, the presence of **D**s in the population, both in the Lower-Degree and Medium-Degree classes, renders some High-Degree individuals unable to join the rest of the "hubs", remaining with a low wealth. The survival of **D**s is therefore detrimental to the overall wealth of the population and, individually, **D**s fare pretty badly in strongly heterogeneous communities, down to small values of the enhancement factor r. Diversity provides indeed a powerful mechanism to promote cooperation.

Dependence on population size and average connectivity

In the main text the discussion relied on simulations carried out for fixed population size N and fixed average degree z. Here we investigate how the results depend on these 2 quantities. To this end we carried out simulations for $N=500,1000$ and 5000 (for fixed $z=4$) and $z=4,16,32,64$ for fixed $N=500$. The results for fixed $z=4$ are shown in Figures S2 and S3, whereas results for fixed N and varying z are shown in Figure S4.

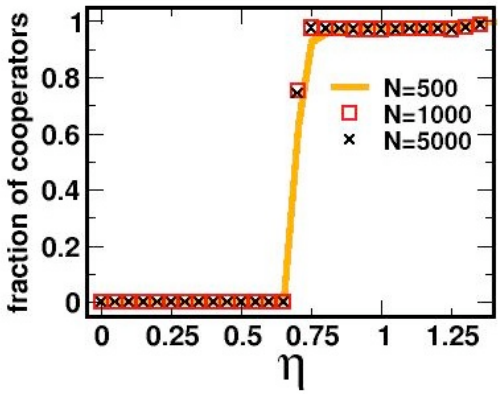

Figure S2. **Dependence of the evolution of cooperation on population size – regular graphs.** Results show that the evolution of cooperation is independent of population size for fixed average connectivity z. Populations are modeled in terms of regular graphs and each point corresponds to an average over 2000 runs for populations of sizes $N=500,1000,5000$. Note that, for regular graphs, the results do not depend on the cost paradigm: fixed cost per game, or fixed cost per individual.

Figures S2 and S3 evidence the negligible dependence of our results on the overall population size. We have checked that, for small z, the results are valid down to population sizes of $N\approx250$, below which the results are less smooth, given the increasing role of finite-size fluctuations on the evolutionary dynamics.

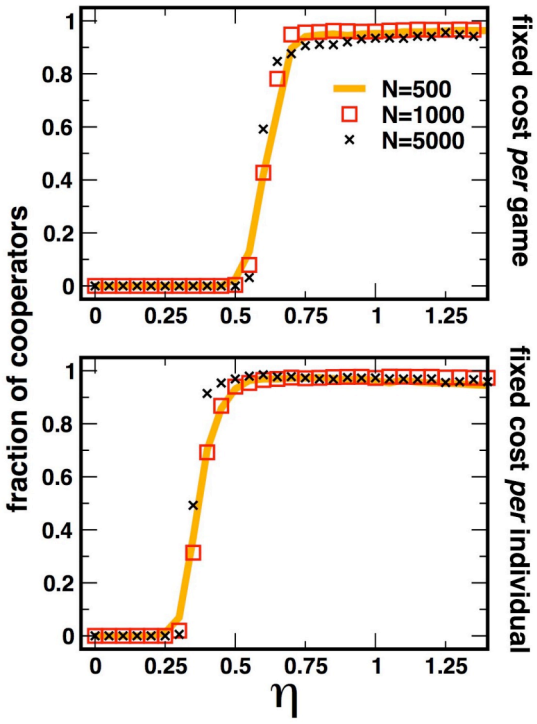

Figure S3. **Dependence of the evolution of cooperation on population size – Scale-free graphs.** Following the same notation of Figure S2, the results show that the evolution of cooperation is independent of population size for fixed average connectivity.

Concerning the dependence of the results on the average connectivity z (the number of links of the graph for fixed population size), several factors work against cooperative behavior when the average connectivity *z* increases.

First, the average size of the groups *(z+1)* increases. This factor induces an overall scaling of the value of the multiplication factor *r*: Since in Figure S4 we plot the fraction of cooperators as a function of η, this scaling is automatically included, which means that remaining differences between curves are due to other factors. The results of Figure S4 show that, irrespective of the cost paradigm, the critical value of η above which cooperators no longer get extinct does not qualitatively change with the average connectivity and is equivalent to an overall rescaling of *r*. This reflects the important role played by the average group size *(z+1)*.

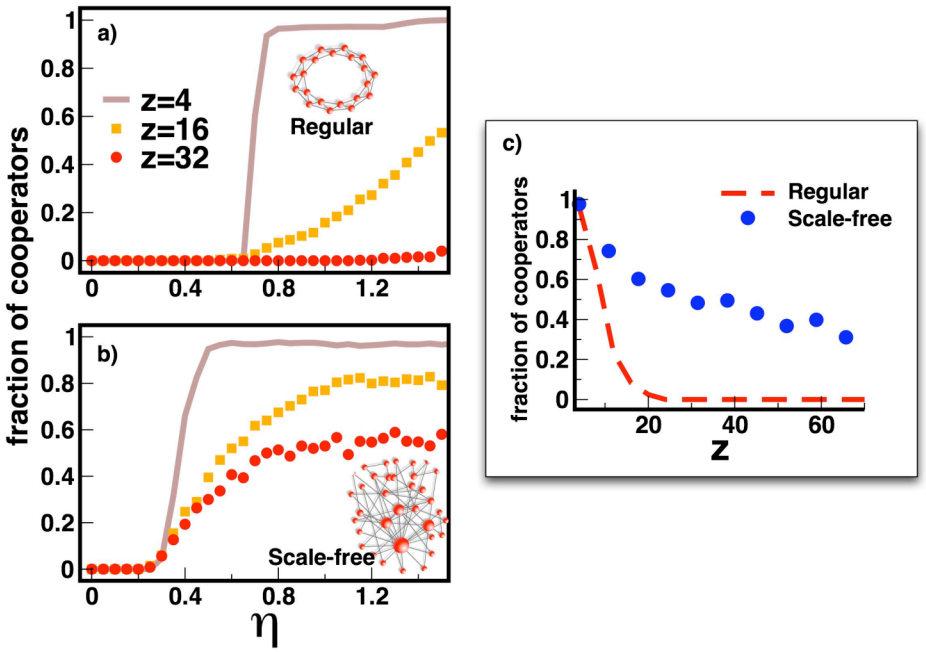

Figure S4. **Evolution of cooperation under PGG in populations with different average connectivity z.** Figures a) and b) follow the same notation and scheme of Figures 3 and 4. In c) we plot the final fraction of cooperators (for fixed η=0.8 in a region of coexistence of defectors and cooperators – see Figure S3). With increasing average degree z, results for PGG on graphs follow a trend similar to that obtained for simple two-player games on graphs[16] as one would expect (see main text for details). As z becomes sizeable cooperation will inevitably collapse as average group size increases and the overall degree of heterogeneity (on scale-free graphs) also decreases. Each value corresponds to an average over 100 runs for 20 network realizations and populations with N=500 individuals.

Furthermore, changing z towards a fully-connected graph corresponds to the limit case of having a PGG with the size of the whole (finite) population. In this limit, the presence of a single defector will always lead to the demise of cooperation, independently of the value of r. For instance, Figure S4 shows that, on regular networks, $z = 16$ is enough to reduce significantly the final number of cooperators, even when η>1. Similar to simple 2-player games, spatially constrained populations are only able to sustain sizeable levels of cooperative behavior on sparse graphs[8, 18, 23]

– lack of diversity leaves survival of cooperation contingent on the feasibility of cooperators to form tight communities. With increasing connectivity, these tight communities become increasingly vulnerable to exploitation, favouring defectors.

Figure S4-c also shows that the significant boost of cooperation obtained for scale-free graphs compared to regular graphs is a robust feature, being a slowly decreasing function of the average connectivity z. The results of Figure S4 correspond to the situation of a fixed investment per individual, but the conclusions remain qualitatively independent of the investment scheme adopted.

Cooperators and Defectors on the star(s)

In the scale-free graphs we have used to model heterogeneous populations, individuals have never less than 2 neighbors. In Fig. S5-a we provide a detail of such a graph. This feature determines the occurrence of short closed loops (for example triangles, which are responsible for the small, yet non-zero value of the cluster coefficient exhibited by these graphs) that precludes simple analytic treatments.

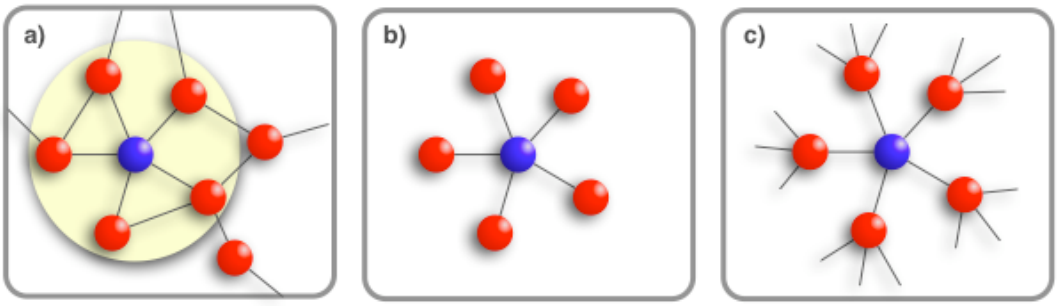

Figure S5. **The star and generalized star graphs.** We shall employ the simple star graph depicted in b), in this case 1 center and 5 leaves, as the simplest abstraction of the sub-graph selected in panel a). In panel c) we generalize the star such that every leaf has $k-1$ links, the overall structure exhibiting no loops.

In order to demonstrate the mechanism responsible for the emergence of cooperation in Figure 2-b, we resort here to the simplest possible (and most disadvantageous) situation for a cooperator – that of a single **C** in a population of **D**s only. Moreover, we shall start by "abstracting" from the connections of the neighbors of this **C** to other **D**s, which naturally leads to a star-graph, Figure S5-b. To the extent that the single **C** may have a larger fitness than any of her **D**-neighbors, the **C**-strategy will spread. The fact that on a star there are only 2 types of nodes - center and leaves - and no loops clearly simplifies the mathematical analysis.

Let us then consider a star of size N : 1 center (h) and N-1 leaves (l) - with the single **C** located on the center (Figure S5-b). In such scenario (if the **C** is placed on a leaf her fitness will never exceed that of a **D** in the center), the fitness (Π) of the **C** is given by

$$\Pi(h) = \frac{rc}{N^2} + (N-1)\frac{r}{2}\frac{c}{N} - c$$

whereas the fitness of any of the **D**s (leaves) reads

$$\Pi(l) = \frac{rc}{N^2} + \frac{rc}{2N} .$$

From the expressions above, the **C**-strategy will spread whenever

$$\Pi(h) - \Pi(l) > 0 \Leftrightarrow r > \frac{2}{1-2/N}$$

a *decreasing* function of N. In other words, not only is it possible for a single **C** to become advantageous, but also the critical value of the multiplication factor r decreases with increasing number of leaves. This result suggests that the feasibility of the **C**–strategy to spread increases with the connectivity of the node on which the **C** is located.

Clearly, the star is a gross simplification of a realistic population structure; a useful, yet simple, structure, is the generalized star depicted in Figure S5-c, in which every leaf has *k-1* external links to other **D**-neighbors. In this case we obtain

$$r > \frac{k}{1-2/N}$$

which remains a decreasing function of N (for fixed k). Clearly, invasion is easier the larger the difference (diversity) between the connectivities of the center and of the leaves. On the other hand, whenever $k=N$, we obtain the limit of a "homogeneous generalized star", in which we may write for the critical threshold

$$r > \frac{N}{1-2/N}.$$

On scale-free populations, the majority of individuals have few connections, whereas a few highly connected individuals ensure the overall connectivity of the entire population. Naturally, most of the highly connected individuals (hubs) connect to individuals of low connectivity, which provide excellent conditions for a **C** located on a hub to spread to other nodes.

So far we have seen that **C**s can spread more efficiently from centers to leaves on heterogeneous graphs. But how do **C**s invade a **D** hub ? Similar to the star-model considered above, the double star model below provides a simple illustration of the evolutionary dynamics of "hub-invasion".

Let us consider the generalized double-star configuration depicted in Figure S6.

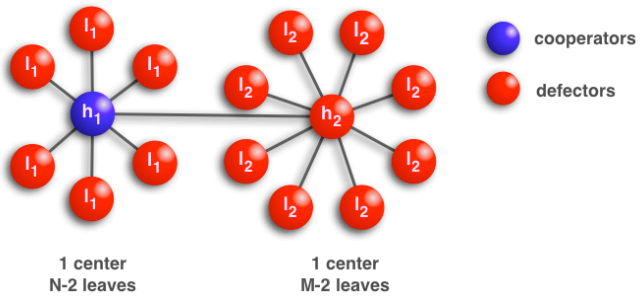

Figure S6. **The double-star graph.** In the double star depicted, we have 2 centers (h_1 and h_2) the left one with *N-2* leaves and the right one with *M-2* leaves. This simple (heterogeneous) structure allows us to inspect the conditions under which a single cooperator in h_1 may take over a defector in h_2. We shall in fact consider a generalized double-star, in which case each l_1 has *k-2* external links (no loops) whereas each l_2 has *j-2* external links.

We have 2 centers (hubs h_1 and h_2) with *N-2* and *M-2* leaves respectively, and 1 link connecting the two centers. Each of the *N-2* leaves has additional *k-2* external links, whereas each of the *M-2* leaves has *j-2* external links. Let us place a **C** on h_1 with the rest of the nodes as **D**s. In this case, the **C**-strategy will spread to h_2 whenever

$$\Pi(h_1) - \Pi(h_2) > 0 \Leftrightarrow r > \frac{kN}{N-2} \equiv \alpha \,,$$

whereas the **C**-strategy will spread to the leaves of h_1 whenever

$$\Pi(h_1) - \Pi(l_1) > 0 \Leftrightarrow r > \frac{kMN}{k + M(N-3)} \equiv \beta \ .$$

Since $\alpha < \beta$ (whenever $M > k$), it will be easier for the **C** sitting on h_1 to invade h_2 than the leaves of h_1. In fact, invasion of h_2 depends only on the number of leaves of h_1 and of their connectivity: the larger the leaf-connectivity the harder it will be to invade h_2, whereas the larger the connectivity of h_1, the easier it will be to invade h_2. In particular, notice that the invasion condition does not depend on the connectivity of h_2 ! In other words, even on "generalized" double-stars, **C**s will manage to expand to the extent that $r > \beta$.

In the remainder of this section, we would like to use the simple double-star to show how **C**s resist invasion by **D**s. This mechanism is crucial to the survival of cooperation[27]. To this end we introduce a **D** in h_1 in a population of all **C**s, as shown in Figure S7-a.

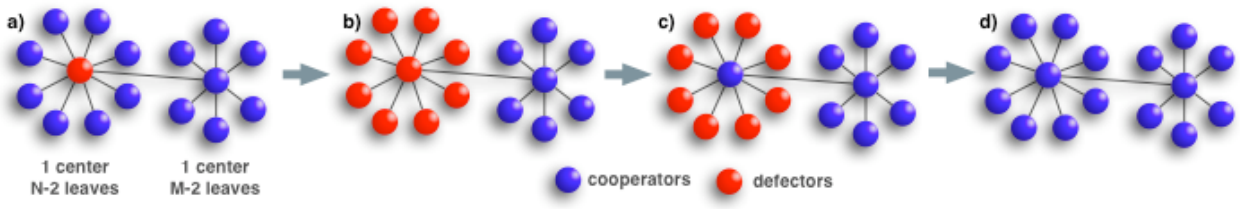

Figure S7. **The demise of a successful D.** Panel a) shows a single **D** in a sea of **C**s. Whenever the **D**-fitness is larger than the fitness of any of the **C**s, the **D**-strategy will spread. We show that such spreading occurs preferentially to the leaves, which contributes to reduce the fitness of the center **D**, making it vulnerable to a take over by the central **C**. Such negative feedback mechanism of the **D**s leads to their own demise.

Whether the single **D** is the fittest individual in the double star will depend on the balance between N and M. Indeed, we obtain

$$\Pi(h_2) > \Pi(h_1) \Leftrightarrow r > \frac{4M}{M^2 - 4 - M(N-2)}$$

and to the extent that r satisfies the inequality above, the **C**-center will have a larger fitness than the **D**-center and consequently will spread the **C**-strategy onto the **D**-star. Another interesting condition is the one equating the fitness of both centers if, instead of a neighborhood with **C**s only, the **D**-center has only $k \leq N-2$ **C** neighbors. We may write

$$\Pi(h_2) > \Pi(h_1^{\,k}) \Leftrightarrow r > \frac{4M}{M^2 - 4 - Mk}.$$

Let us imagine, however, that we start with *k=N-2* and (for example) r is such that $\Pi(h_2) < \Pi(h_1)$. In this case the **D** strategy will spread. However, unlike the symmetric situation discussed before, here we get

$$\left[\Pi(h_1) - \Pi(l_1)\right] - \left[\Pi(h_1) - \Pi(h_2)\right] = \frac{rc}{4M^2}\left[4 + M\left(M^2 + M - 8\right)\right] > 0 \ (M > 2)$$

Consequently, it is more likely that **D** will spread to l_1 than to h_2, given that each player will imitate her neighbor with a probability proportional to the payoff difference. This means the neighborhood of h_1 will turn into **D** before h_2, creating the pattern of Figure S7-b (in practice, it need not reach the configuration in Figure S7-b as whenever $\Pi(h_2) > \Pi(h_1^k)$ the **C**-center may actually invade the **D**-center). In other words, **D**s are victims of their own success, as they efficiently spread their strategy to the "weak" neighbors, reducing their own fitness and becoming prone to be taken over by the **C**-center.

At this stage (Figure S7-b), the payoffs of both centers are

$$\Pi(h_1) = \frac{rc}{MN} + \frac{r}{M}\left(\frac{(M-2)c}{2} + \frac{c}{M}\right)$$

$$\Pi(h_2) = \Pi(h_1) + rc\frac{M^2 - 4}{4M} - c$$

such that h_2 easily becomes advantageous with respect to h_1 (M>2)

$$\Pi(h_2) > \Pi(h_1) \Leftrightarrow r > \frac{4M}{M^2 - 4}.$$

In other words, evolutionary dynamics leads the population into the configuration of Figure S7-c (or, even better, the **D**-center will be taken over by the **C**-center before all **C**s on the leaves of the left turn into **D**s). At this stage we return to a configuration similar to the one of Figure S6, but more beneficial for **C**s: The spreading of the **C**-strategy to the leaves on the left will take place whenever ($\{M, N\} > 2$)

$$r > \frac{2M^2N}{2(M+N)+2MN(M-1)-3M^2}$$

which clearly favors invasion of the leaves, even for small N and M. In practice, simulation results show (Figure S8) that **C**s will tend to dominate the population.

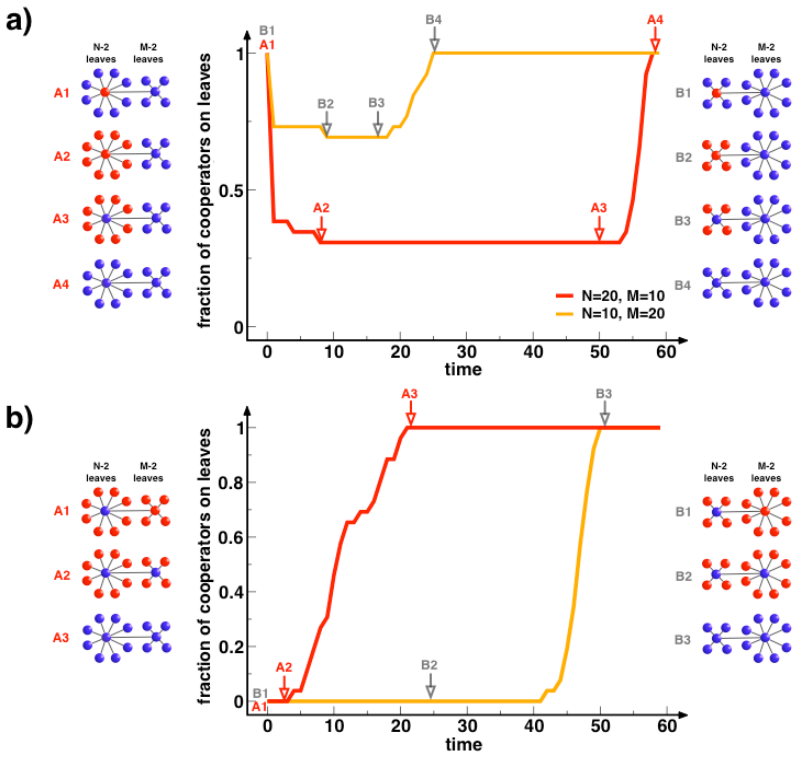

Figure S8. **Evolution of cooperation under PGG on double stars.** Figures a) and b) provide a typical scenario for time evolution of the fraction of **C**s on the leaves of double-star graphs. **a)** We start with a single **D** both in the largest center (red line) and in the smallest center (orange line) (configurations A1 and B1). The overall behavior shows that the **D** invades the leaves of her star (A2 and B2), after which is invaded by the **C** in the center of the second star (A3 and B3). Subsequently, the remaining defectors on the leaves are invaded by **C**s (B4) (see Figure S7). As expected, the relative connectivity of both centers (**C** and **D**) determines the overall time required for invasion (red line: r=1.3, N=20 and M=10 ; orange line: r=1.3, N=10 and M=20). **b)**, We start with a single **C** located in one of the centers in a population of **D**s (A1 and B1), with r=2.8 (r>β>α, see main text). The **C**-center starts by invading the **D**-center (A2 and B2), after which the C-strategy spreads to all leaves (A3 and B3). Given that the ability of the **C**-center to invade the **D**-center increases with the **C**-center connectivity, $N>M$ leads to a faster invasion than $N<M$.

Finally, it is worth noting that the introduction of small loops in the double star above leads to important changes in the overall evolutionary dynamics; not only the time for **C**-dominance decreases, but also the critical values of r above which **C**s dominate also decreases, a feature which is difficult to capture analytically.