Climate policies under wealth inequality

Vítor V. Vasconcelos[a,b,c], Francisco C. Santos[a,c], Jorge M. Pacheco[a,d,e], and Simon A. Levin[f,g,h,1]

[a]Applications of Theoretical Physics Group, Centro de Matemática e Aplicações Fundamentais, Instituto para a Investigação Interdisciplinar, P-1649-003 Lisbon Codex, Portugal; [b]Centro de Física da Universidade do Minho, 4710-057 Braga, Portugal; [c]Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento (INESC-ID) and Instituto Superior Técnico, Universidade de Lisboa, Taguspark, 2744-016 Porto Salvo, Portugal; [d]Departamento de Matemática e Aplicações, Universidade do Minho, 4710-057 Braga, Portugal; [e]Centro de Biologia Molecular e Ambiental (CBMA), Universidade do Minho, 4710-057 Braga, Portugal; [f]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003; [g]University Fellow, Resources for the Future, Washington, DC 20036; and [h]Fellow, Beijer Institute of Ecological Economics, SE-104 05 Stockholm, Sweden

Contributed by Simon A. Levin, December 18, 2013 (sent for review October 3, 2013)

Taming the planet's climate requires cooperation. Previous failures to reach consensus in climate summits have been attributed, among other factors, to conflicting policies between rich and poor countries, which disagree on the implementation of mitigation measures. Here we implement wealth inequality in a threshold public goods dilemma of cooperation in which players also face the risk of potential future losses. We consider a population exhibiting an asymmetric distribution of rich and poor players that reflects the present-day status of nations and study the behavioral interplay between rich and poor in time, regarding their willingness to cooperate. Individuals are also allowed to exhibit a variable degree of homophily, which acts to limit those that constitute one's sphere of influence. Under the premises of our model, and in the absence of homophily, comparison between scenarios with wealth inequality and without wealth inequality shows that the former leads to more global cooperation than the latter. Furthermore, we find that the rich generally contribute more than the poor and will often compensate for the lower contribution of the latter. Contributions from the poor, which are crucial to overcome the climate change dilemma, are shown to be very sensitive to homophily, which, if prevalent, can lead to a collapse of their overall contribution. In such cases, however, we also find that obstinate cooperative behavior by a few poor may largely compensate for homophilic behavior.

collective action | global warming | governance of the commons | environmental agreements | evolutionary game theory

Despite existing scientific consensus that anthropogenic greenhouse gas emissions (GHGE) perturb global climate patterns with negative consequences for many natural ecosystems (1–3), reaching a global agreement regarding reduction of GHGE remains one of the most challenging problems humans face (4). International climate negotiations have largely failed to reach consensus (5, 6), evidencing a conflict between rich and poor countries, which often do not agree on the urgency of emission reduction measures, given the scientific uncertainty regarding the impacts of climate change (7–10). Indeed, in the aftermath of the 15th Conference of Parties in Copenhagen/2009 one has observed a tendency of several governments to regard climate change as a problem of a distant future—2050—hence discounting (4) the actual risk of collective disaster—despite predictions that severe climate change consequences, such as increased occurrence of heat waves and droughts, for instance, may happen sooner (1).

The issue of reducing GHGE has been addressed recently, both experimentally and theoretically, by means of a threshold Public Goods Game (PGG) in which success requires overall cooperative collective action, and decisions must be made knowing that failure to cooperate implies a risk of overall collapse (10–18). Like many social dilemmas of collective action, any participant that curbs emissions pays a cost whereas the benefits are shared among everyone. Thus, the rational choice is to free ride on the benefits produced by others at their own expense (through abatement), leading to the well-known tragedy of the commons, where selfish behavior results in overexploitation of the public good (19, 20).

Both theory and experiment agree that risk perception plays a central role in escaping the tragedy of the commons (12, 13). Besides risk, the role of wealth inequality among participants has been recently investigated by means of economic experiments involving students from western, educated, industrialized, rich, and democratic (WEIRD) countries (a feature that may induce biases regarding behavior of subjects taking the role of poor countries) (10, 11). Games comprised groups of fixed size ($N = 6$) where participation was equally split between rich and poor individuals, whose different wealth resulted from two different start-up amounts of money made available to group participants. The insights provided by these experiments (10, 11) (using different methodologies and assumptions while using the same PGG) converge on the idea that resolution of the climate change policy problem stems from the rich compensating for the smaller contribution by the poor and, even when risk is very high (something that does not seem to apply to the present situation), there is still a very significant chance of failing to solve the climate change dilemma, a situation that is ameliorated whenever intermediate tasks are designated (11) or whenever individuals have the opportunity to pledge their contribution before actual action (10).

Here we address the issue of wealth inequality from a theoretical perspective. The model we extend here to deal with wealth inequality (13, 17) has been shown to lead to predictions that correlate nicely with previous economic experiments carried out in the absence of any wealth inequalities (10), with the added value of allowing a full exploration of how success in addressing the climate change dilemma depends on other important parameters, such as risk, group size, introduction of sanctioning institutions of global or local nature, etc. It is important to stress that,

Significance

One of the greatest challenges in addressing global environmental problems such as climate change, which involves public goods and common-pool resources, is achieving cooperation among peoples. There are great disparities in wealth among nations, and this heterogeneity can make agreements much more difficult to achieve (e.g., regarding implementation of climate change mitigation). This paper incorporates wealth inequality into a public goods dilemma, including an asymmetric distribution of wealth representative of existing inequalities among nations. Without homophily (imitation of like agents), inequality actually makes cooperation easier to achieve; homophily, however, can undercut this, leading to collapse because poor agents may contribute less. Understanding such effects may enhance the ability to achieve agreements on climate change and other issues.

Author contributions: V.V.V., F.C.S., J.M.P., and S.A.L. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: slevin@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323479111/-/DCSupplemental.

despite the limited number of scenarios realizable in the laboratory, data stemming from behavioral experiments have provided crucial insights, not only because they unravel human behavior when confronted with climate change issues, but also because they provide important guidelines that help in calibrating theoretical models, such as the one we use here. Indeed, here we model climate change negotiations in the framework of Evolutionary Game Theory, where individuals exhibit a well-defined behavior (C or D), as a result of which they accumulate a certain payoff resulting from the game group interactions. Regularly, every individual A compares her/his payoff with that of a randomly chosen individual B, imitating (or not) the behavior of B with a probability that is a growing function of the payoff difference between B and A. Naturally, the larger the payoff of individual B (randomly) selected as role model, the more likely it is that A imitates her/his behavior (see *Methods* for specific details of the update rule).

Let us consider a population of finite size Z of which Z_R individuals are rich and $Z_P = Z - Z_R$ individuals are poor. Individuals are randomly sampled from the population and organize into groups of size N. Each rich individual starts with an initial endowment b_R whereas each "poor" individual starts with b_P, with $b_R > b_P$. These endowments may be used (or not) by an individual to contribute to reducing GHGE in her own group. We distinguish two types of behavior: (*i*) cooperators (Cs), who contribute a certain fraction c of their endowment to help solve the group task, and (*ii*) defectors (Ds) that do not contribute anything to solve the group task. Hence, the endowment is directly related to what each participant will lose if the next intermediate target is not met: Cs will lose $b_{R/P}(1-c)$ whereas Ds will lose the entire endowment $b_{R/P}$. If the overall amount of contributions in the group is above a certain threshold $Mc\overline{b}$ (where $\overline{b}$ is the average endowment of the population), the target will be met. Otherwise, with a probability r—the perception of risk of collective disaster (10, 12, 13)—individuals in the group will lose whatever they have.

This framework creates an interdependent behavioral ecosystem, where each player in a group knows what all other members of the group will do and where decisions and achievements of others influence one's own decisions (21–24). In particular, decisions taken by the poor can be potentially influenced by the actions and achievements of the rich (and vice versa), adding an additional coupling between these two subpopulations (details in *Methods*). In standard conditions, anyone in this population may influence and be influenced by anyone else. This, however, may not always be the case, in the sense that individuals may be more receptive to the behavior and decisions of those in the same wealth class, thus selecting preferentially those of their wealth class as peers. To this end we define a homophily parameter ($0 \le h \le 1$), such that when $h = 1$, individuals are restricted to influence (and be influenced) by those of the same wealth status, whereas when $h = 0$, no wealth discrimination takes place. Naturally, such influence dynamics occur in the presence of action errors (24) as well as other stochastic effects, such as random exploration of the strategy space, akin to behavioral mutations (25).

In *SI Text*, we show that in populations with a mixed composition of rich and poor (and for different combinations of Cs and Ds in each wealth group), the nature of the overall public goods dilemma faced by the rich subpopulation differs qualitatively from that faced by the poor subpopulation. In these limiting cases where a decoupling of the timescales associated with the dynamics of the rich and of the poor takes place, the rich generally face an N-person coordination dilemma between Cs and Ds, whereas (for most combinations of parameters) poor Cs and poor Ds engage in a coexistence dilemma. Such a diversity in the nature of the games played—due to, e.g., heterogeneous interaction patterns or resource distributions—will have strong implications on the emerging social dynamics, often promoting the chances of achieving cooperation in structured populations (13, 26, 27).

In practice, however, no reason other than mathematical simplicity may justify such extreme scenarios. When analyzing the fully coupled dynamics on the entire configuration space represented by a two-dimensional simplex (see, e.g., Fig. 2), in which the y axis (x axis) portrays the fraction of Cs among the poor (rich), it is natural to ask to which extent the existence of rich and poor alters the dynamics of the risky PGG at stake, compared with the standard model where no wealth inequality is explicitly considered. To answer this question, we start by recognizing that 20% of the world's wealthier countries produce approximately the same gross domestic product as the remaining 80%. Thus, we break the population into two wealth classes, such that the poor comprise 80% of the population, whereas the rich constitute the remaining 20%. Concomitantly, we assume that poor countries contribute an amount proportional to their wealth (as reflected in their initial endowment) and similarly with the rich. As a result, different groups will exhibit, on average, different ratios of rich and poor, reflecting the intrinsic wealth asymmetry that one observes in the real world.

In Fig. 1 we compare the average group achievement (η_G), that is, the fraction of time a group succeeds in achieving $Mc\overline{b}$ as a function of risk (*Methods*), in the cases when there is no wealth inequality (gray line) and in the presence of wealth inequality (blue and red lines). The results show unequivocally that wealth inequality may promote group success. This result, however, depends strongly on the level of homophily (h): Whenever the rich and poor are evenly influenced by anyone else (no homophily, $h = 0$, blue line), group achievement is enhanced for all values of risk (r). However, when the rich (poor) influence and are influenced by rich (poor) only (homophily $h = 1$, red line), the chances of success are generally below those attained in the absence of wealth inequality.

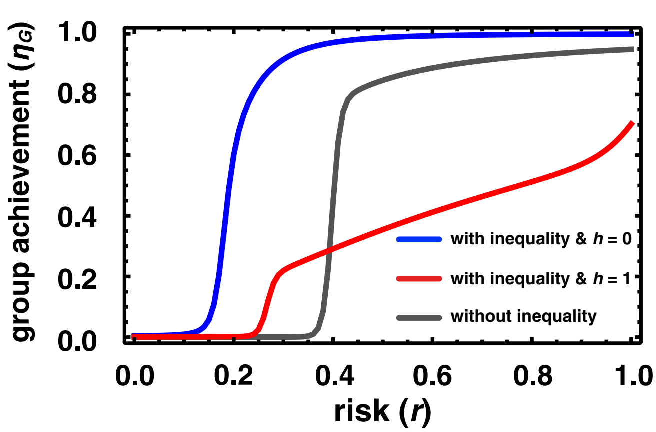

Fig. 1. Average group achievement η_G as a function of risk. The gray line shows the average group achievement in the case of no wealth inequality; that is, all individuals have an initial endowment $b = \overline{b} = 1$ and the cost of cooperation is $0.1b$. The blue line shows results for wealth inequality with the homophily parameter $h = 0$, whereas the red line shows results for $h = 1$. We split the population of $Z = 200$ individuals into $Z_R = 40$ rich (20%) and $Z_P = 160$ poor (80%); initial endowments are $b_R = 2.5$ and $b_P = 0.625$, ensuring that the average endowment $\overline{b}$ remains $\overline{b} = 1$ (used to generate the gray line); the cost of cooperation also remains, on average, $0.1\overline{b}$, which means $c_R = 0.1b_R$ and $c_P = 0.1b_P$. The results show that wealth inequality significantly enhances the average chance of group success in the absence of homophily ($h = 0$), whereas under homophily ($h = 1$) the fact that only like influences like brings the overall chances of success to levels generally below those under wealth equality. Other parameters (*Methods*) are $N = 6$ and $M = 3$.

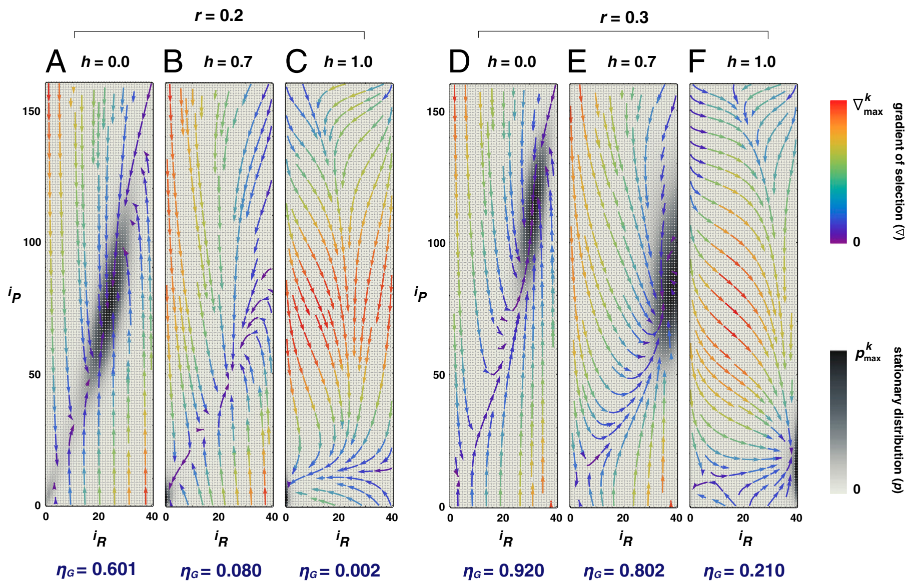

Fig. 2. Stationary distribution and gradient of selection for different values of risk r and of the homophily parameter h. (*A–F*) Each panel contains all possible configurations of the population (in total $Z_R \times Z_P$), each specified by the number of rich (i_R) and poor (i_P) it contains and represented by a gray-colored dot. Darker dots represent those configurations in which the population spends more time, thus providing a contour representation of the stationary distribution. The curved arrows show the so-called gradient of selection (∇), which provides the most likely direction of evolution from a given configuration. We use a color code in which red lines are associated with higher speed of transitions. The behavioral dynamics of the population depend on the homophily parameter h in a nonlinear way. For $h \leq 0.5$, the results remain qualitatively similar to those depicted for $h = 0$, in which case everybody influences and is influenced by everybody else. In this case, the contribution of the rich is sizeable, which also leads the poor to contribute. For $h > 0.5$ the behavior changes abruptly, and one witnesses the rapid collapse of cooperation among the poor and, for low risk ($r = 0.2$, *A–C*), an ensuing disappearance of contributions to the overall PGG, with the population spending most of the time in full defection, leading to a dramatic impact on the overall group achievement η_G, indicated below each contour plot. However, a slight increase in overall risk perception (here $r = 0.3$, *D–F*) actually impels the rich to contribute, despite the fact that the poor still do not cooperate. Other parameters: $Z = 200$; $Z_R = 40$; $Z_P = 160$; $c = 0.1$; $N = 6$; $M = 3c\bar{b}$ ($\bar{b} = 1$); $b_P = 0.625$; $b_R = 2.5$; $p^k_{max} = \{p^A_{max}, \cdots, p^F_{max}\} = \{2, 40, 75, 3, 2, 20\} \times 10^{-3}$; and $\nabla^k_{max} = \{\nabla^A_{max}, \cdots, \nabla^F_{max}\} = \{16, 6, 2, 16, 6, 3\} \times 10^{-2}$.

Fig. 1 also shows that η_G depends strongly on homophily, a feature that plays a central role in the overall dynamics, as is discussed in more detail below. However, it is also important to understand how the contributions are split between the rich and the poor, a feature that is not possible to grasp directly from η_G. To this end we now study in detail the stationary distributions associated with the dynamics of the two-subpopulation model. The results are shown in Fig. 2. Each (discrete) configuration is represented by a small circle, colored in gray tones. Darker circles indicate those configurations visited more often, providing a representation of the full stationary distribution ($\bar{p}$, *Methods*), i.e., the prevalence in time of each possible configuration of the population. Arrows in each simplex represent, in turn, the most probable direction of evolution starting from a given configuration (the gradient of selection ∇, *Methods*). For each arrow, we adopt a continuous color code associated with the likelihood of such a transition (brighter colors indicate more likely transitions).

Fig. 2 shows that, even in the absence of significant homophily bias ($h \leq 0.5$) a higher fraction of rich contribute (with average values of 57% for $r = 0.2$ and 78%, for $r = 0.3$), compared with the poor (with average values of 46% for $r = 0.2$ and 69% for $r = 0.3$), thus also protecting their greater wealth. This result does not depend on risk; however, for low risk, the overall contribution is limited, increasing significantly after a slight increase in overall risk perception. Indeed, in the absence of homophily, cooperation may prevail in a wealth-unequal world (e.g., Fig. 2*D*).

Qualitatively, one can now understand the results in Fig. 1 if one takes into account that, in most cases, the dynamics both among the rich and among the poor can become dominated by basins of attraction that lead to a coexistence between *C*s and *D*s (*SI Text*). Whenever the risk is moderate to high, there is an increase of the size of such basins, with a corresponding increase of the stationary fraction of cooperators, such that the feedback dynamics between the poor and the rich act to build up the cooperation levels among both subpopulations. In other words, the poor pave the way for the rich to cooperate, which, in turn, feeds back into the poor, also increasing their levels of cooperation. This feedback occurs because, perhaps counterintuitively, not only the poor imitate the rich, but also the rich imitate the poor. In fact, it is easy to prove that, for the model considered, the rich imitate the poor more often than the poor imitate the rich.

As also shown in Fig. 2, this positive feedback between the two subpopulations is interrupted whenever homophily becomes dominant ($h \sim 1$). When rich and poor cease to be able to sway one another, we observe two distinct scenarios: At low risk ($r = 0.2$ in Fig. 2) overall cooperation collapses. With a slight increase in risk perception, however ($r = 0.3$), the rich contribute, despite the fact that the poor do not. Together with risk, a lack of homophily plays an important role: As soon as the homophily constraint is relaxed—by adopting $h < 1$—poor individuals start to be nudged by the successes of the rich, effectively inducing the poor players to contribute to the common good.

However, even in the absence of homophily ($h = 0$), this positive feedback between the two subpopulations does not always lead to an increase of cooperation—thus we obtain the coexistence dynamics shown in Fig. 2. Indeed, whenever most poor opt for cooperation, the dynamics drive rich countries toward less cooperation, given that they may now profit from the larger overall contributions stemming from the poor. Similar dynamics may also occur among the poor. This reduction, however, not only does not prevent the majority of rich from engaging in cooperation, but also does not compromise the overall group achievement values. As a result of these coupled dynamics, the population will stay most of the time nearby a coexistence equilibrium (interior attractor, Fig. 2 *A*, *D*, and *E*).

This said, we are all aware that some individuals may be more receptive than others to change their mind, based on the influence of their peers. In fact, some individuals—for various reasons, as witnessed in the world summits on climate change that have taken place to date—may maintain the same behavior irrespective of their sphere of influence. Given the small size of the overall population, such an obstinate behavior may lead to sizeable effects in the global dynamics. In the following we investigate how such obstinate behaviors (in both wealth classes) affect the overall dynamics. For simplicity, we assume that, in all cases, obstinate behavior amounts to 10% of individuals in one subpopulation—which corresponds to the same fixed contribution to the PGG, considering either rich or poor obstinate players—see *SI Text* for a more detailed analysis.

Fig. 3 shows that obstinate poor cooperators provide impressive improvements in the aggregate propensity of the population to achieve coordination ($\eta_G = 0.581$ compared with $\eta_G = 0.004$ in the absence of obstinate individuals), larger than obstinate rich cooperators, who lead to less pronounced enhancements ($\eta_G = 0.223$). This effect, which extends qualitatively to all values of h, is more pronounced when $h = 1$, as is the case in Fig. 3.

The trend shown in Fig. 3 is qualitatively inverted in the case of obstinate defectors who, in general, are detrimental to overall cooperation and group achievement (details in *SI Text*). These results are largely independent of the parameters chosen and highlight the important role that obstinate defectors among the rich and obstinate cooperators among the poor may play in the outcome of climate negotiations.

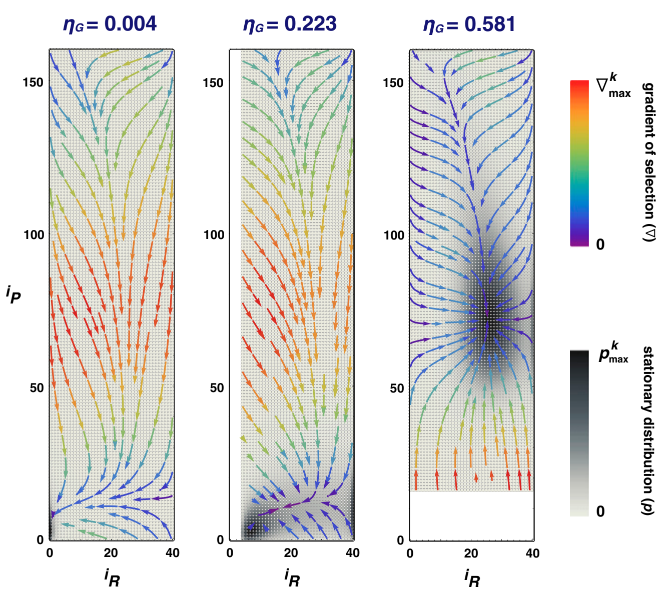

Fig. 3. Stationary distribution and gradient of selection for populations comprising 10% of individuals exhibiting an obstinate cooperative behavior. Same notation as in Fig. 2 is used. Whenever 10% of individuals exhibit obstinate cooperative behavior (*Center* and *Right* contours), the number of configurations of the population in which the evolutionary dynamics proceed is correspondingly reduced (white areas in contours). The *Left* contour contains no obstinate individuals and is displayed for reference. In the *Center* contour, 10% of the rich individuals behave as obstinate cooperators; that is, they never change their behavior. In the *Right* contour, 10% of poor individuals exhibit such behavior. A small fraction of obstinate rich and obstinate poor cooperators lead to very different outcomes, also for the average group achievement η_G. Indeed, the chances of success are significantly enhanced whenever obstinate cooperator behavior occurs among the poor. The effect is most pronounced whenever individuals are homophilic, as is the case here ($h = 1$). Other parameters: $Z = 200$; $Z_R = 40$; $Z_P = 160$; $c = 0.1$; $N = 6$; $M = 3c\bar{b}$ ($\bar{b} = 1$); $b_P = 0.625$; $b_R = 2.5$; $r = 0.2$; $\beta = 5.0$; $p^k_{\max} = \{p^A_{\max},\ p^B_{\max},\ p^C_{\max}\} = \{76,\ 4,\ 2\} \times 10^{-3}$; and $\nabla^k_{\max} = \{\nabla^A_{\max},\ \nabla^B_{\max},\ \nabla^C_{\max}\} = \{3,\ 3,\ 4\} \times 10^{-2}$.

In summary, homophily generally impels the rich to compensate for the poor. Given that contributions from the poor are crucial to solving the climate change problem we face, it is then imperative that homophilic behavior is avoided (it is noteworthy that it is enough that overall behavior is not purely homophilic for homophily to be effectively avoided). Moreover, a small fraction of obstinate poor cooperators leads to sizable increases in the overall prospects for success, mostly when homophily rules. Perhaps unsurprisingly, when the contribution of the poor is widespread, the rich refrain from contributing. Certainly, David Hume would not be impressed by this feature that emerges from the game dynamics.

Conventional wisdom would lead one to believe that wealth inequality and homophily would constitute important obstacles regarding overall cooperation in climate change negotiations. Our results predict that, as long as (*i*) risk perception is high; (*ii*) climate negotiations are partitioned in smaller groups agreeing on local, short-term targets; and (*iii*) individuals are influenced by their more successful peers, whom they imitate—irrespective of their wealth class—and making errors while doing so, the prospects are not that grim. On the contrary we find that, under such conditions, cooperation may outcompete defection, benefiting from wealth inequality. Thus, hope remains that the problem may be overcome. Moreover, the qualitative nature of the results obtained here remains robust if we assume that, instead of proportional contributions, poor and rich contribute the same amount, when cooperating.

Our model, however, ignores an important factor: that the thresholds may be intrinsically uncertain. This uncertainty, if sizeable, can destroy cooperation, as sharply demonstrated recently, both theoretically and experimentally (9). Likely, to the extent that agreements aim at short-term targets involving smaller groups, it will also be easier to narrow down threshold uncertainties. Nonetheless, and in the absence of wealth inequality, introducing threshold uncertainty into our model leads to the same scenarios predicted (and confirmed) in ref. 9 (*SI Text*).

Finally, the recent report of the Intergovernmental Panel for Climate Change (28), besides emphasizing that climate change is real and humans are the main cause of it, urging countries to stop the warming of the planet, has also attempted to narrow down the threshold uncertainty. However, given that risk perception is low and that a bottom–up approach [as defended by the late Elinor Ostrom (29) and also, indirectly, by the results of the present model] has yet to spread globally, it is perhaps not surprising that today's prospects remain gloomy. Clearly it is urgent that individuals become aware of the true risk that we face. Indeed, an increase in risk perception will surely promote the development of local initiatives that may foster overall cooperation by extending the bottom–up approach to all players of the global game.

Methods

We consider a population of Z individuals, Z_R of which are considered rich (initial endowment b_R) and Z_P considered poor (initial endowment b_P) who, together, set up groups of size N, in which they engage in the climate change threshold PGG (12, 13). Each individual is capable of adopting one of the two strategies: *C* and *D*. Following the discussion in the main text, and given that rich *C*s contribute $c_R = cb_R$ whereas poor *C*s contribute $c_P = cb_P$, the payoff of an individual playing in a group in which there are j_R rich *C*s, j_P poor *C*s, and $N - j_R - j_C$ *D*s, can be written as $\Pi^D_{R/P} = b_{R/P}\{\Theta(\Delta) + (1-r)[1-\Theta(\Delta)]\}$ and $\Pi^C_{R/P} = \Pi^D_{R/P} - c_{R/P}$ ($\Delta = c_R j_R + c_P j_P - Mc\bar{b}$), for rich/poor *D*s and *C*s, respectively. In the equations above, $\Theta(k)$ is the Heaviside function [that is, $\Theta(k) = 1$ whenever $k \geq 0$, being zero otherwise], $0 < M \leq N$ is a positive integer, $\bar{b}$ is the average endowment ($Z\,\bar{b} = Z_R b_R + Z_P b_P$), and r (the perception of risk) is a real parameter varying between 0 and 1; the parameters $c < 1$, $\bar{b}$, b_R, and b_P are all positive real numbers. Finally, the fitness f^X of an individual adopting a given strategy, X, will be associated with the average payoff of that strategy in the population. The average payoff can be computed for a given strategy in a configuration $\mathbf{i}=\{i_R,i_P\}$, using a multivariate hypergeometric sampling (without replacement) (details in *SI Text*). The number of individuals adopting a given strategy will evolve in time according to a stochastic birth–death process combined with the pairwise comparison rule (24, 30), which describes the social dynamics of rich *C*s, poor *C*s, rich *D*s, and poor *D*s in a finite population. Under pairwise comparison, each individual of strategy X adopts the strategy Y of another member of the population, with probability given by the Fermi function $(1+e^{\beta(f^X-f^Y)})^{-1}$, where β controls the intensity of selection ($\beta = 3$ in Figs. 1 and 2, $\beta = 5$ in Fig. 3). In the absence of homophily, the strategy Y is chosen at random with uniform probability. For a finite value of the homophily parameter h, individuals of the same wealth class are chosen with probability 1 whereas individuals of the other wealth class are chosen with probability $1 - h$; thus, when homophily is maximum, the choice occurs only among the individuals of the same wealth class (rich or poor) (details in *SI Text*). Additionally, we consider that, with a mutation probability μ ($\mu = 1/Z$ in Figs. 1–3), individuals adopt a randomly chosen strategy. As the evolution of the system depends only on its actual configuration, evolutionary dynamics can be described as a Markov process over a two-dimensional space. Its probability distribution function, $p_\mathbf{i}(t)$, which provides information on the prevalence of each configuration at time t, obeys a master equation (details in *SI Text*), a gain–loss equation involving the transition rates between all accessible configurations (24, 31, 32). The stationary distribution $\bar{p}_\mathbf{i}$ is then obtained by reducing the master equation to an eigenvector search problem (31) (details in *SI Text*). Another central quantity that portrays the overall evolutionary dynamics in the space of all possible configurations is the gradient of selection $\nabla_\mathbf{i}$. For each configuration $\mathbf{i}$, we compute the most likely path the population will follow, resorting to the probability to increase (decrease) the number of individuals adopting a strategy S_k, $T_\mathbf{i}^{S_k+}$ ($T_\mathbf{i}^{S_k-}$) in each time step. Additionally, for each possible configuration $\mathbf{i}$, we make use of multivariate hypergeometric sampling to compute the (average) fraction of groups that reach a total of $Mc\bar{b}$ in contributions, that is, that successfully achieve the public good—which we designate by $a_G(\mathbf{i})$. Average group achievement—η_G—is then computed, averaging over all possible configurations $\mathbf{i}$, each weighted with the corresponding stationary distribution $\eta_G=\sum_\mathbf{i}\bar{p}_\mathbf{i}a_G(\mathbf{i})$.

ACKNOWLEDGMENTS. The authors thank Flávio L. Pinheiro for useful discussions. J.M.P. thanks all members of the Department of Ecology and Evolutionary Biology, Princeton University, for their support during his stay. This research was supported by Fundação para a Ciência e Tecnologia (FCT) through Grants SFRH/BD/86465/2012, PTDC/MAT/122897/2010, and EXPL/EEI-SII/2556/2013; by multiannual funding of CBMA and INESC-ID (under Projects PEst-C/BIA/UI4050/2011 and PEst-OE/EEI/LA0021/2013) provided by FCT, and by Fundação Calouste Gulbenkian through the "Stimulus to Research" program for young researchers. S.A.L. acknowledges financial support from National Science Foundation Grant EF-1137894.

1. Meehl GA, et al. (2007) Global climate projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds Solomon S, et al. (Cambridge Univ Press, New York), pp 747–845.
2. Alley RB, et al. (2007) Summary for policy makers. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ Press, New York).
3. Stott PA, et al. (2010) Detection and attribution of climate change: A regional perspective. *Wiley Interdiscip Rev Clim Change* 1(2):192–211.
4. Levin SA (2012) The trouble of discounting tomorrow. *Solutions* 4(3):20–24.
5. Barrett S (2005) *Environment and Statecraft: The Strategy of Environmental Treaty-Making* (Oxford Univ Press, New York).
6. Barrett S (2007) *Why Cooperate? The Incentive to Supply Global Public Goods* (Oxford Univ Press, New York).
7. O'Neill BC, Oppenheimer M (2002) Climate change. Dangerous climate impacts and the Kyoto Protocol. *Science* 296(5575):1971–1972.
8. Oppenheimer M, Alley RB (2005) Ice sheets, global warming, and Article 2 of the UNFCCC. *Clim Change* 68(3):257–267.
9. Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. *Proc Natl Acad Sci USA* 109(43):17372–17376.
10. Tavoni A, Dannenberg A, Kallis G, Löschel A (2011) Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proc Natl Acad Sci USA* 108(29):11825–11829.
11. Milinski M, Röhl T, Marotzke J (2011) Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Clim Change* 109:807–814.
12. Milinski M, Sommerfeld RD, Krambeck HJ, Reed FA, Marotzke J (2008) The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc Natl Acad Sci USA* 105(7):2291–2294.
13. Santos FC, Pacheco JM (2011) Risk of collective failure provides an escape from the tragedy of the commons. *Proc Natl Acad Sci USA* 108(26):10421–10425.
14. Wang J, Fu F, Wu T, Wang L (2009) Emergence of social cooperation in threshold public goods games with collective risk. *Phys Rev E Stat Nonlin Soft Matter Phys* 80(1 Pt 2):016101.
15. Abou Chakra M, Traulsen A (2012) Evolutionary dynamics of strategic behavior in a collective-risk dilemma. *PLoS Comput Biol* 8(8):e1002652.
16. Chen X, Szolnoki A, Perc M (2012) Risk-driven migration and the collective-risk social dilemma. *Phys Rev E Stat Nonlin Soft Matter Phys* 86(3 Pt 2):036101.
17. Vasconcelos VV, Santos FC, Pacheco JM (2013) A bottom-up institutional approach to cooperative governance of risky commons. *Nature Clim. Change* 3(9):797–801.
18. Abou Chakra M, Traulsen A (2014) Under high stakes and uncertainty the rich should lend the poor a helping hand. *J Theor Biol* 341:123–130.
19. Hardin G (1968) The tragedy of the commons. *Science* 162(3859):1243–1248.
20. Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ Press, Cambridge, UK).
21. Fowler JH, Christakis NA (2010) Cooperative behavior cascades in human social networks. *Proc Natl Acad Sci USA* 107(12):5334–5338.
22. Carman KG (2003) Social influences and the private provision of public goods: Evidence from charitable contributions in the workplace. Discussion paper 02–13 (Stanford Institute for Economic Policy Research, Stanford, CA).
23. Sigmund K (2010) *The Calculus of Selfishness* (Princeton Univ Press, Princeton).
24. Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E Stat Nonlin Soft Matter Phys* 74(1 Pt 1):011909.
25. Traulsen A, Hauert C, De Silva H, Nowak MA, Sigmund K (2009) Exploration dynamics in evolutionary games. *Proc Natl Acad Sci USA* 106(3):709–712.
26. Santos FC, Santos MD, Pacheco JM (2008) Social diversity promotes the emergence of cooperation in public goods games. *Nature* 454(7201):213–216.
27. Kun Á, Dieckmann U (2013) Resource heterogeneity can facilitate cooperation. *Nat Commun* 4:2453.
28. Intergovernmental Panel on Climate Change (2013) *Climate Change 2013: The Physical Science Basis*. Available at http://www.ipcc.ch/report/ar5/wg1/#.Us3i7fRdV8G. Accessed January 9, 2014.
29. Ostrom E (2009) A polycentric approach for coping with climate change. Policy Research Working Paper, Background Paper to the 2010 World Development Report, WPS5095. Available at http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-5095. Accessed January 9, 2014.
30. Szabó G, Toke C (1998) Evolutionary prisoner's dilemma game on a square lattice. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 58(1):69–73.
31. Kampen NV (2007) *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam).
32. Imhof LA, Fudenberg D, Nowak MA (2005) Evolutionary cycles of cooperation and defection. *Proc Natl Acad Sci USA* 102(31):10797–10800.

Supporting Information

Vasconcelos et al. 10.1073/pnas.1323479111

SI Text

Theoretical and Experimental Studies of Behavior Regarding Climate Change

Several experiments have been performed to understand human behavior in dealing with global warming (1–4). Social dilemmas involving collective action were set up in a repeated game framework, where a given threshold had to be surpassed—otherwise, there was a variable risk (externally defined) of everyone losing all their endowments. The first results have shown that, most of the time, disaster was not avoided (1), risk being an important factor in promoting disaster avoidance. Later on the possibility to make pledges was introduced in the same repeated threshold public goods game, showing that pledges led to an increase of cooperation (despite the possibility of acting differently from what was pledged) (2). However, when the same treatments were run using players with different wealth, this improvement was demoted (2). Using the same game settings, two time horizons were introduced into the dilemma (3), showing that fixing intermediate goals in climate agreements is beneficial, although the final target is reached less often than the intermediate target. More recently, a nonrepeated threshold public goods game experiment was performed in which the effects of both impact uncertainty (where individuals could lose a random amount of money if the threshold was not met) and threshold uncertainty (where the threshold was a random value) were investigated (4). The authors set up the experiment based on results from a theoretical analysis of such a game (assuming fully rational individuals engaging in a one-shot threshold public goods game) that predicted the existence of a critical value for the threshold uncertainty above which cooperation would collapse. Experiments fully confirmed this catastrophic prediction.

Despite the limited number of scenarios realizable in the laboratory, data stemming from behavioral experiments have provided crucial insights, not only because they unravel human behavior when confronted with climate change issues, but also because they provide important guidelines toward developing theoretical models (5–9). Indeed, a dynamical approach to the problem of cooperating to tame the planet's climate was developed, inspired by the intriguing results that experiments were revealing (5–7), allowing one not only to generalize the experimental settings to scenarios that are more difficult to realize in the laboratory, but also to predict what the impact of different approaches to the solution of the climate change problem may bring. Namely, the effect of risk perception and the disruptive power of uncertainty have been captured in the models (*SI Text*, *Threshold Uncertainty*). Theoretical models also extended the experimental insights by predicting the importance of small groups and stringent requirements in improving cooperation, as well as the role and scale of sanctioning institutions in supervising agreements. In keeping with this discussion, the present model brings additional information to this important subject.

In particular, although conventional wisdom would lead one to believe that wealth inequality and homophily would constitute important obstacles regarding overall cooperation in climate change negotiations, the present model predicts that, as long as (*i*) risk perception is high; (*ii*) climate negotiations are partitioned into smaller groups agreeing on local, short-term targets; and (*iii*) individuals are influenced by their more successful peers, whom they imitate—irrespectively of their wealth class—and making errors while doing so, the prospects are not that grim. On the contrary we find that, under such conditions, cooperation may outcompete defection, benefiting from wealth inequality. On the other hand, and to the extent that agreements aim at short-term targets involving smaller groups, it may also become easier to narrow down threshold uncertainties that, if large, do haunt overall cooperation (4) (*SI Text*, *Threshold Uncertainty*).

Evolutionary Dynamics in Finite Populations Under Wealth Inequality, Uncertainty, and Homophily

Let us consider a population of Z individuals. As stated in the main text, each individual adopts one of the two possible strategies $X \in \{C, D\}$ and belongs to one of two possible wealth classes $k \in \{R, P\}$. Let us assume there are Z_R rich (with initial endowment b_R) and Z_P poor individuals (with an initial endowment b_P). These numbers will remain fixed. Individuals are given an initial endowment (with $b_P < b_R$) and play the climate threshold Public Goods Game (PGG) (1, 5), engaging in groups of size N. Following the discussion in the main text, and given that rich Cs contribute $c_R = cb_R$ whereas poor Cs contribute $c_P = cb_P$, the payoff of an individual playing in a group in which there are j_R rich Cs, j_P poor Cs, and $N - j_R - j_P$ Ds can be written as

$$\Pi^D_{R/P}(j_R, j_P) = b_{R/P}\{\Theta(c_R j_R + c_P j_P - Mc\bar{b}) + (1-r)[1 - \Theta(c_R j_R + c_P j_P - Mc\bar{b})]\}$$

and $\Pi^C_{R/P}(j_R, j_P) = \Pi^D_{R/P}(j_R, j_P) - c_{R/P}$, for rich/poor Cs and Ds, respectively. In the equations above, $\Theta(k)$ is the Heaviside function [that is, $\Theta(k) = 1$ whenever $k \geq 0$, being zero otherwise], $0 < M \leq N$ is a positive integer, $\bar{b}$ is the average endowment ($Z\,\bar{b} = Z_R b_R + Z_P b_P$), and r (the perception of risk) is a real parameter varying between 0 and 1; the parameters $0 < c < 1$, $\bar{b}$, b_R, and b_P are all real positive. Finally, the fitness f^X_k of an individual adopting a given strategy X in a population of wealth class k will be associated with the average payoff of that strategy in the entire population. This can be computed for a given configuration of strategies and wealth classes specified by $\mathbf{i} = \{i_R, i_P\}$, using a multivariate hypergeometric sampling (without replacement):

$$f^C_R(\mathbf{i}) = \binom{Z-1}{N-1}^{-1} \sum_{j_R=0}^{N-1} \sum_{j_P=0}^{N-1-j_R} \binom{i_R - 1}{j_R}\binom{i_P}{j_P} \times \binom{Z - i_R - i_P}{N-1-j_R-j_P} \Pi^C_R(j_R + 1, j_P) \quad \text{[S1a]}$$

$$f^D_R(\mathbf{i}) = \binom{Z-1}{N-1}^{-1} \sum_{j_R=0}^{N-1} \sum_{j_P=0}^{N-1-j_R} \binom{i_R}{j_R}\binom{i_P}{j_P} \times \binom{Z - 1 - i_R - i_P}{N-1-j_R-j_P} \Pi^D_R(j_R, j_P) \quad \text{[S1b]}$$

$$f^C_P(\mathbf{i}) = \binom{Z-1}{N-1}^{-1} \sum_{j_R=0}^{N-1} \sum_{j_P=0}^{N-1-j_R} \binom{i_R}{j_R}\binom{i_P - 1}{j_P} \times \binom{Z - i_R - i_P}{N-1-j_R-j_P} \Pi^C_P(j_R, j_P + 1) \quad \text{[S1c]}$$

$$f^D_P(\mathbf{i}) = \binom{Z-1}{N-1}^{-1} \sum_{j_R=0}^{N-1} \sum_{j_P=0}^{N-1-j_R} \binom{i_R}{j_R}\binom{i_P}{j_P} \times \binom{Z - 1 - i_R - i_P}{N-1-j_R-j_P} \Pi^D_P(j_R, j_P). \quad \text{[S1d]}$$

The number of individuals adopting a given strategy will evolve in time according to a stochastic birth–death process combined with the pairwise comparison rule (10, 11), which describes the social dynamics of rich *C*s, poor *C*s, rich *D*s, and poor *D*s in a finite population. Under pairwise comparison, each individual of strategy X adopts the strategy Y of a randomly selected member of the population, with probability given by the Fermi function $(1+e^{\beta(f_k^X-f_l^Y)})^{-1}$, for any wealth class $\{k,l\}\in\{R,P\}$, where β controls the intensity of selection. Additionally we consider that, with a mutation probability μ, individuals adopt a randomly chosen different strategy, in such a way that when $\mu=1$, the individual does change strategy. As the evolution of the system depends only on its actual configuration, evolutionary dynamics can be described as a Markov process over a two-dimensional space. Its probability distribution function, $p_{\mathbf{i}}(t)$, which provides information on the prevalence of each configuration at time t, obeys a master equation of the form

$$p_{\mathbf{i}}(t+\tau)-p_{\mathbf{i}}(t)=\sum_{\mathbf{i}'}\{T_{\mathbf{i}\mathbf{i}'}p_{\mathbf{i}'}(t)-T_{\mathbf{i}'\mathbf{i}}p_{\mathbf{i}}(t)\}, \qquad \textbf{[S2]}$$

a gain–loss equation that allows one to compute the evolution of $p_{\mathbf{i}}(t)$ given the transition probabilities per unit time τ from the configuration $\mathbf{i}$ to $\mathbf{i}'$, $T_{\mathbf{i}'\mathbf{i}}$ (11–13). The stationary distribution $\bar{p}_{\mathbf{i}}$ analyzed in the main text is obtained by making the left-hand side equal to zero, which transforms Eq. **S2** into an eigenvector search problem (12), namely, the eigenvector associated with the eigenvalue 1 of the transition matrix W (12) whose matrix elements W_{qp} are built from the transition probabilities per unit time $T_{\mathbf{i}'\mathbf{i}}$ in the following way: Let us enumerate each of all possible configurations $\mathbf{i}=\{i_R,i_P\}$ of the population by an integer number—we do so by defining a bijective function V such that $p=V(\mathbf{i})$ and $q=V(\mathbf{i}')$ and, therefore, $\mathbf{i}=V^{-1}(p)$ and $\mathbf{i}'=V^{-1}(q)$. Then, we may write $W_{qp}=T_{\mathbf{i}'\mathbf{i}}$. The transition probabilities $T_{\mathbf{i}'\mathbf{i}}$ can all be written in terms of the following expression, which gives the probability that an individual with strategy $X\in\{C,D\}$ in the subpopulation $k\in\{R,P\}$ changes to a different strategy $Y\in\{C,D\}$, both from the same subpopulation k and from the other population l (that is, $l=P$ if $k=R$, and $l=R$ if $k=P$):

$$T_k^{X\to Y}=\frac{i_k^X}{Z}\left((1-\mu)\left[\frac{i_k^Y}{Z_k-1+(1-h)Z_l}\left(1+e^{\beta\left(f_k^X-f_k^Y\right)}\right)^{-1}+\frac{(1-h)i_l^Y}{Z_k-1+(1-h)Z_l}\left(1+e^{\beta\left(f_k^X-f_l^Y\right)}\right)^{-1}\right]+\mu\right).$$

Thus, if the homophily is maximum ($h=1$), the imitation occurs only between individuals of the same wealth class (rich or poor), whereas $h=0$ means that everyone influences and may be influenced by anyone else.

Another central quantity—which portrays the overall evolutionary dynamics in the space of all possible configurations—is the gradient of selection $\nabla_{\mathbf{i}}$ (GoS). For each configuration $\mathbf{i}=\{i_R,i_P\}$, we compute the most likely path each subpopulation $k\in\{R,P\}$ will follow, resorting to the probability to increase (decrease) by one, in each time step, the number of cooperators for that configuration $\mathbf{i}$ of the population, which we denote by $T_{\mathbf{i},k}^{+}$ ($T_{\mathbf{i},k}^{-}$), such that

$$\nabla_{\mathbf{i}}=\left\{T_{\mathbf{i},R}^{+}-T_{\mathbf{i},R}^{-},\ \ T_{\mathbf{i},P}^{+}-T_{\mathbf{i},P}^{-}\right\}.$$

Finally, for each possible configuration $\mathbf{i}$, we make use of multivariate hypergeometric sampling (Eq. **S1**) to compute the (average) fraction of groups that reach a total of $Mc\bar{b}$ in contributions, that is, that successfully achieve the public good—which we designate by $a_G(\mathbf{i})$. Average group achievement—η_G—is then computed by averaging over all possible configurations $\mathbf{i}$, each weighted with the corresponding stationary distribution $\eta_G=\sum_{\mathbf{i}}\bar{p}_{\mathbf{i}}a_G(\mathbf{i})$.

Timescale Separation: Games Among the Rich and Among the Poor

To assess what games the rich play in the presence of the poor (*C*s and *D*s) and the poor play in the presence of the rich (*C*s and *D*s), we let each subpopulation evolve assuming that the rate of evolution of the other subpopulation is zero. The results are shown in Fig. S1, where we compute the gradient of selection (∇) that governs the evolutionary dynamics of the rich in the presence of frozen, mixed configurations of the poor (Fig. S1 *A* and *B*) and vice versa (Fig. S1 *C* and *D*). We consider the cases in which the population is subdivided into subpopulations of equal size ($Z_P=Z_R$, Fig. S1 *A* and *C*) or not ($Z_P=4Z_R$, as in the main text, Fig. S1 *B* and *D*).

The results in Fig. S1 show that the rich tend to be more cooperative as the difference between the endowments of both classes increases, whereas cooperation among the poor largely remains unaffected. Moreover, whereas the poor engage in a coexistence game in which overall cooperation decreases as cooperation among rich decreases, the dynamics of the rich are influenced by the relative size of the poor subpopulation. In general, the rich engage in an *N*-player stag-hunt game (14) with different degrees of coordination and coexistence, depending on the (fixed) fraction of poor cooperators. As a result, different combinations of parameters may transform the original *N*-player stag-hunt dilemma (14)—characterized by two internal roots—into a pure coordination or coexistence dilemma or even a defection dominance dilemma (Fig. S1 *A* and *C*). However, as discussed before (5, 6) and illustrated in Fig. 2 of the main text, the unstable fixed point can be overcome by stochastic effects—such as errors in imitation and random exploration of the strategy space—such that the population spends most of its time in the vicinity of the coexistence points. Thus, the prevalent levels of cooperation among the rich will be ultimately defined by the size of the cooperative basin attraction and the position of the respective coexistence root, both influenced by the dynamics occurring among the poor. It is also noteworthy that the gradients of the rich are 10 times smaller than those of the poor. This means the rate of response of the rich to changes is (on average) slower than that of the poor. In practice, the poor will adjust their behavior more rapidly to changes of the configuration of the rich, thus quickly shifting between the corresponding levels of coexistence between poor *C*s and *D*s.

Evolutionary Dynamics for the Same Amounts of Rich and Poor

In all experimental settings carried out to date, the fraction of rich and poor in each group was kept equal. Here we compute the analog situation in our model; that is, we compute the stationary distribution in the case when $Z_P=Z_R$. The results are shown in Fig. S2.

Comparison with Fig. 2 in the main text shows unequivocally that, for low risk ($r=0.2$), rich and poor populations of the same size lead to more pessimistic prospects concerning overall cooperation (compare the values of η_G below each panel). For a corresponding increase of risk perception (Fig. 2), we observe that, overall, cooperation remains below that observed for asymmetric subpopulation sizes. In particular, the rich cooperate less and are no longer able to compensate for the collapse of cooperation among the poor, a feature that becomes more pronounced with increasing homophily.

Threshold Uncertainty

In the absence of wealth inequality, all individuals are equivalent. This has been, to date, the most studied situation in the laboratory. In particular, in ref. 4 it has been demonstrated how, in a situation that all individuals in the group are equivalent, threshold uncertainty has a disruptive effect on the overall chances of cooperation. Below we show that this is also the case in our model.

Therefore, we modify the individual payoffs so that the games played have a random threshold, with a value drawn from a uniform probability distribution in the interval $[Mc\overline{b}-\delta, Mc\overline{b}+\delta]$. The larger the value of δ is, the larger the uncertainty associated with the threshold. Fig. S3*A* shows how this uncertainty induces a regime shift in the overall behavior of the population, from an *N*-player coordination game (5, 14, 15) where cooperators do have a chance toward a defection dominance dilemma. This shift leads, in turn, to a radical change in the profile of the stationary distribution, also shown in Fig. S3*B*. The impact of this threshold uncertainty on group achievement, $\eta_G(r)$, is shown in Fig. S4, corroborating the results obtained in ref. 4. The study of the effects of threshold uncertainty in the presence of wealth inequality, which is more complex given the problem that the threshold does not affect in the same way the rich and the poor, will be deferred to a future study.

Evolutionary Dynamics in the Presence of Obstinate Cooperators and Defectors

Here we investigate in more detail the role played by obstinate individual behavior in the population. We assume that a fixed fraction of individuals in the population exhibits obstinate behavior; that is, these individuals are not susceptible to changing their behavior in time. We compute, for all possible combinations, the overall group achievement η_G (*Methods* and *SI Text*, *Evolutionary Dynamics in Finite Populations Under Wealth Inequality, Uncertainty, and Homophily*) shown in Fig. S5 for three different values of risk (0.2, 0.3, and 0.4, each associated with a different line color) and for the fractions of obstinate individuals indicated in Fig. S5 *A*–*G*, *Insets*.

We carry out this analysis as a function of the homophily parameter h. The results corroborate the idea that obstinate *C*s generally lead to positive effects in what concerns overall group achievement, whereas obstinate *D*s lead to negative effects. Among these, obstinate poor *C*s play a crucial role in sustaining cooperation, mostly when homophily is high ($h \sim 1$), whereas obstinate *D*s are generally detrimental to overall cooperation.

Robustness of Results as a Function of *N*

In the following we investigate the dependence of our model results when we change group size and group threshold. To this end we compute, as a function of risk, the same curves that we plot in Fig. 1 of main text, for two group sizes ($N = 6$ and $N = 12$) and for several combinations of M and N, leading to six different scenarios. We use the same parameters as those used in Fig. 1; namely, we split the population of $Z = 200$ individuals into $Z_R =$ 40 rich (20%) and $Z_P = 160$ poor (80%); initial endowments are $b_R = 2.5$ and $b_P = 0.625$, ensuring that the average endowment $\overline{b}$ remains $\overline{b} = 1$ (used to generate the gray line in Fig. 1 and Fig. S6); the cost of cooperation also remains, on average, $0.1\overline{b}$, which means $c_R = 0.1b_R$ and $c_P = 0.1b_P$. The results are shown in Fig. S6. Clearly, group size constitutes a very important parameter, because smaller groups lead to higher chances of success (5). Nonetheless, what we observe, in all cases, is that the message contained in Fig. 1 remains valid for all combinations of parameters shown: Wealth inequality without homophily (blue line) systematically fosters overall cooperation for lower values of risk than what is observed under wealth equality (gray line). Finally, homophilic and wealth-unequal subpopulations lead to the grimmest prospects for overall cooperation.

1. Milinski M, Sommerfeld RD, Krambeck HJ, Reed FA, Marotzke J (2008) The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc Natl Acad Sci USA* 105(7):2291–2294.
2. Tavoni A, Dannenberg A, Kallis G, Löschel A (2011) Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proc Natl Acad Sci USA* 108(29):11825–11829.
3. Milinski M, Röhl T, Marotzke J (2011) Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Clim Change* 109:807–814.
4. Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. *Proc Natl Acad Sci USA* 109(43):17372–17376.
5. Santos FC, Pacheco JM (2011) Risk of collective failure provides an escape from the tragedy of the commons. *Proc Natl Acad Sci USA* 108(26):10421–10425.
6. Santos FC, Vasconcelos VV, Santos MD, Neves P, Pacheco JM (2012) Evolutionary dynamics of climate change under collective-risk dilemmas. *Math Models Methods Appl Sci* 22(Suppl 01):1140004.
7. Vasconcelos VV, Santos FC, Pacheco JM (2013) A bottom-up institutional approach to cooperative governance of risky commons. *Nature Clim Change* 3(9):797–801.
8. Abou Chakra M, Traulsen A (2012) Evolutionary dynamics of strategic behavior in a collective-risk dilemma. *PLoS Comput Biol* 8(8):e1002652.
9. Hilbe C, Abou Chakra M, Altrock PM, Traulsen A (2013) The evolution of strategic timing in collective-risk dilemmas. *PLoS ONE* 8(6):e66490.
10. Szabó G, Toke C (1998) Evolutionary prisoner's dilemma game on a square lattice. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 58(1):69–73.
11. Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E Stat Nonlin Soft Matter Phys* 74(1 Pt 1):011909.
12. Kampen NV (2007) *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam).
13. Imhof LA, Fudenberg D, Nowak MA (2005) Evolutionary cycles of cooperation and defection. *Proc Natl Acad Sci USA* 102(31):10797–10800.
14. Pacheco JM, Santos FC, Souza MO, Skyrms B (2009) Evolutionary dynamics of collective action in N-person stag hunt dilemmas. *Proc Biol Sci* 276(1655):315–321.
15. Souza MO, Pacheco JM, Santos FC (2009) Evolution of cooperation under N-person snowdrift games. *J Theor Biol* 260(4):581–588.

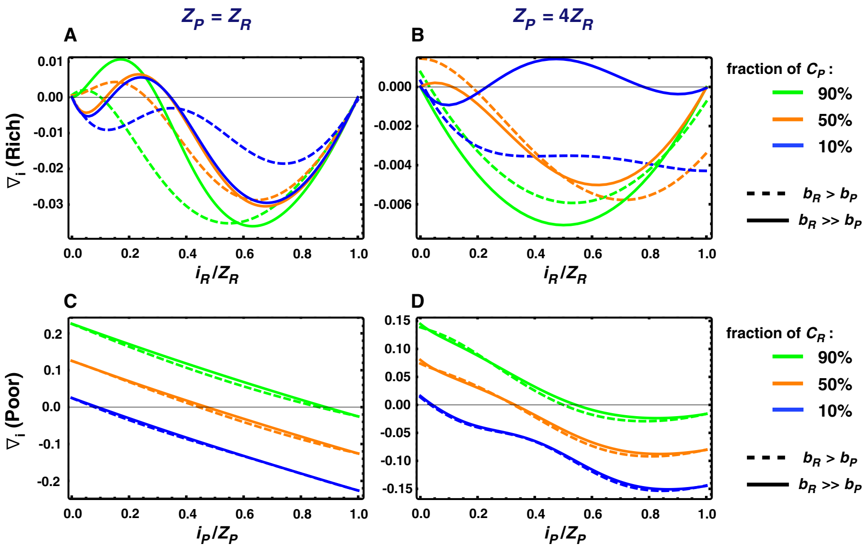

Fig. S1. Rich and poor play different games. (*A*–*D*) The gradient of selection (∇, *Methods*) in the case when one subpopulation (rich or poor) evolves while the rate of evolution of the other subpopulation is zero, in this way assessing which effective game each subpopulation engages in. (*A* and *B*) Number of poor (Z_P) and the fraction of *C*s (i_P/Z_P) in the poor population are kept constant (at three different values, legend at *Right*) while the fraction of *C*s (i_R/Z_R) and *D*s in the rich population is allowed to evolve. (*C* and *D*) Here we reverse the timescales and evolve the poor subpopulation, freezing the composition of the rich subpopulation. For each case we consider two regimes associated with different ratios for the initial endowments of rich and poor: $b_R > b_P$ ($b_R = 1.35$) and $b_R >> b_P$ ($b_R = 1.75$); imposing that $(b_R Z_R + b_P Z_P)/Z = \bar{b} = 1$ fixes the values of b_P for each case. (*A* and *C*) $Z_P = Z_R$. (*B* and *D*) $Z_P = 4Z_R$. Other model parameters: $Z = 200$, $N = 10$, $M = 3$, $\beta = 10$, $h = 0$, $r = 0.3$, $c_R = 0.1b_R$, and $c_P = 0.1b_P$.

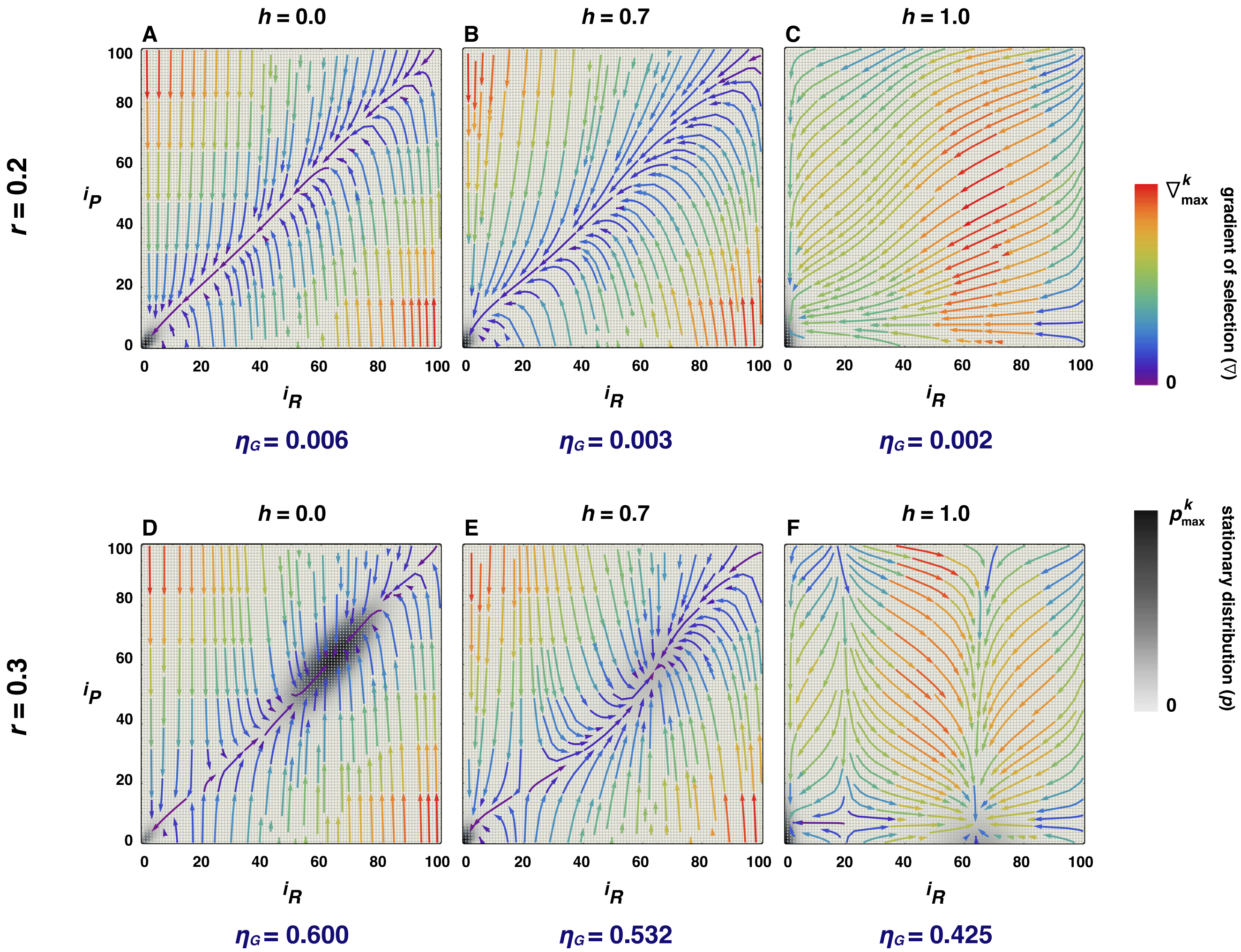

Fig. S2. Evolutionary dynamics for equal fractions of rich and poor. (*A–F*) The gradient of selection (∇, *Methods*) in the case when each subpopulation (rich or poor) evolves with the same number of individuals (compare with Fig. 2). Other model parameters: $Z = 200$, $Z_P = Z_R$, $N = 6$, $M = 3c\overline{b}$, $\beta = 5$, $\mu = 1/Z$, $c_R = 0.1b_R$, $c_P = 0.1b_P$, $b_R = 1.7$, and $b_P = 0.3$, ensuring that $(b_R Z_R + b_P Z_P)/Z = \overline{b} = 1$; $p_{max}^{k=A...F} = \{4.2, 5.3, 3.4, 0.2, 0.7, 1.6\} \times 10^{-2}$; and $\nabla_{max}^{k=A...F} = \{0.25, 0.12, 0.02, 0.25, 0.12, 0.02\}$.

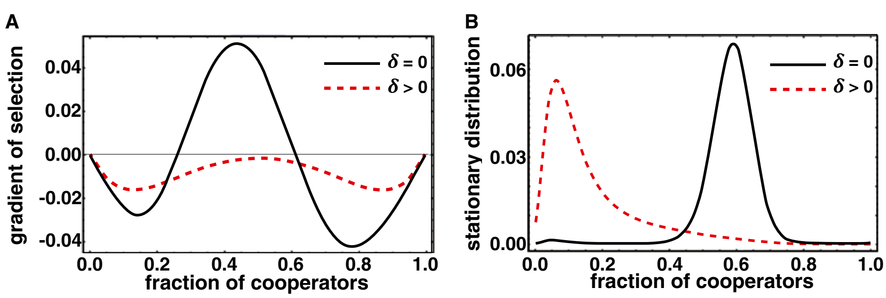

Fig. S3. Threshold uncertainty effect. (*A*) The gradient of selection; (*B*) the stationary distribution, that is, the fraction of time the population spends in each population composition specified by *x*. The black lines provide results for no threshold uncertainty ($\delta = 0$) whereas the red lines show results for $\delta = 2.75$. Other parameters are $Z = 200$, $N = 8$, $M = 4$, $c = 0.1$, $\overline{b} = 1.0$, $\beta = 6.0$, $\mu = 1/Z$, and $r = 0.6$.

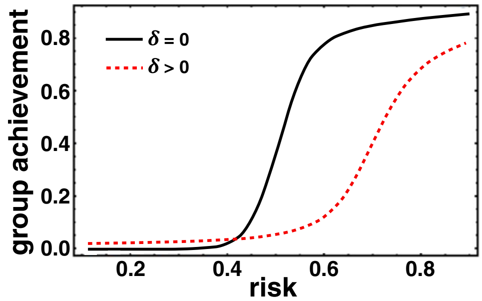

Fig. S4. Group achievement under threshold uncertainty. Results for the group achievement η_G as a function of the risk r are shown. Same parameters as in Fig. S3 are used. Clearly, there is no chance of cooperation before the risk promotes cooperation in an otherwise defection-dominant dilemma imposed by threshold uncertainty.

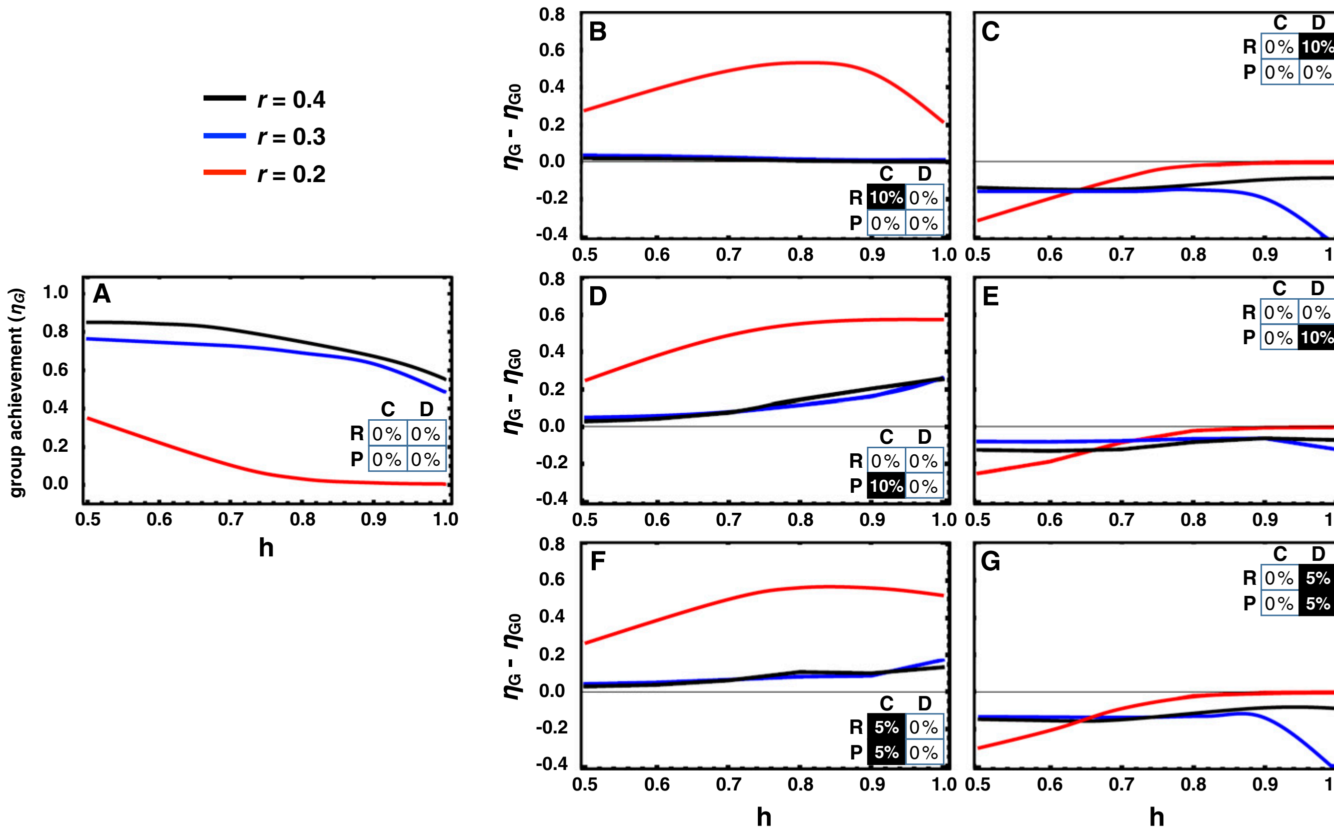

Fig. S5. (*A*–*G*) Average group achievement in the presence of obstinate players. *B*–*G* show the difference between the average group achievement (η_G, *Methods* and *SI Text*, *Evolutionary Dynamics in Finite Populations Under Wealth Inequality, Uncertainty, and Homophily*) in the presence of players whose (obstinate) behavior remains unchanged throughout the evolution of the population and the average group achievement (η_{G0}) computed in the absence of these types of individuals (*A*). Results are shown for the different values of risk indicated. Sizable positive effects (in particular for low risk) are obtained whenever obstinate behavior occurs among cooperators. Other parameters: $Z = 200$, $Z_P = 4Z_R$, $N = 6$, $M = 3$, $\beta = 10$, $\mu = 1/Z$, $c_R = 0.1b_R$, $c_P = 0.1b_P$, $b_R = 2.5$, and $b_P = 0.625$.

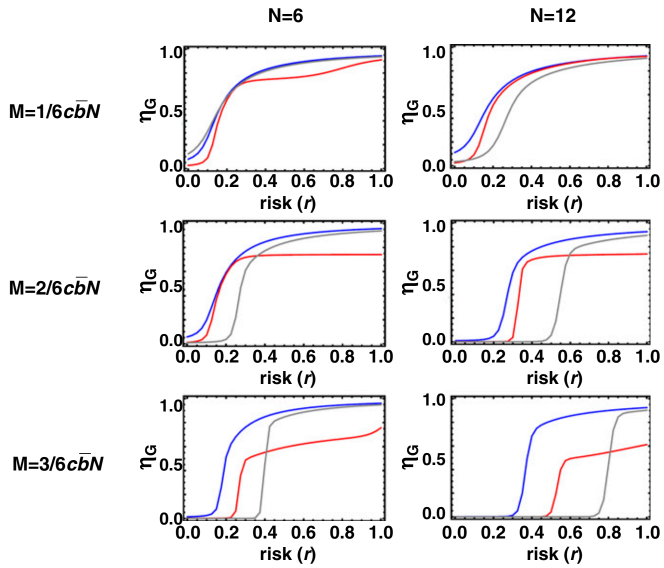

Fig. S6. Robustness of results with respect to *N* and *M*. Different panels show the average group achievement (η_G, *Methods* and *SI Text*, *Evolutionary Dynamics in Finite Populations Under Wealth Inequality, Uncertainty, and Homophily*) as a function of risk *r*, for different group sizes ($N = 6$, *Left*; and $N = 12$, *Right*) and, in each case, for three different values of the sharp threshold *M*: $M = 1$ (*Top*), $M = 2$ (*Middle*), and $M = 3$ (*Bottom*). As group size *N* increases, cooperation blooms for larger values of *r*, a behavior that is more pronounced for larger values of *M*. Nonetheless, and irrespective of the dependences in *N* and *M* exhibited, the overall behavior remains qualitatively identical to that found in Fig. 1 of the main text: Wealth inequality without homophily (blue line) systematically fosters overall cooperation for lower values of risk than what is observed under wealth equality (gray line). Finally, wealth-unequal but homophilic groups lead to the grimmest prospects for overall cooperation. Other parameters: $Z = 200$, $Z_P = 4Z_R$, $\beta = 10$, $\mu = 1/Z$, $c_R = 0.1b_R$, $c_P = 0.1b_P$, $b_R = 2.5$, $b_P = 0.625$.