Diversity enables the jump towards cooperation for the Traveler’s Dilemma

María Alejandra Ramírez1,2*, Matteo Smerlak2, Arne Traulsen1 & Jürgen Jost2,3

Social dilemmas are situations in which collective welfare is at odds with individual gain. One widely studied example, due to the conflict it poses between human behaviour and game theoretic reasoning, is the Traveler’s Dilemma. The dilemma relies on the players’ incentive to undercut their opponent at the expense of losing a collective high payoff. Such individual incentive leads players to a systematic mutual undercutting until the lowest possible payoff is reached, which is the game’s unique Nash equilibrium. However, if players were satisfied with a high payoff -that is not necessarily higher than their opponent’s- they would both be better off individually and collectively. Here, we explain how it is possible to converge to this cooperative high payoff equilibrium. Our analysis focuses on decomposing the dilemma into a local and a global game. We show that players need to escape the local maximisation and jump to the global game, in order to reach the cooperative equilibrium. Using a dynamic approach, based on evolutionary game theory and learning theory models, we find that diversity, understood as the presence of suboptimal strategies, is the general mechanism that enables the jump towards cooperation.

Game theory is a framework for analysing the outcome of the strategic interaction between decision makers1. The fundamental concept is that of a Nash equilibrium where no player can improve her payoff by a unilateral strategy change. Typically, the Nash equilibrium is considered to be the optimal outcome of a game, however in social dilemmas the individual optimal outcome is at odds with the collective optimal outcome2. This means that one player can improve her payoff at the expense of the other by unilaterally deviating, but if both deviate, they end up with lower payoffs. In this type of games, the mutually beneficial, but non-Nash equilibrium strategy is called cooperation. However, in this context cooperation should not be interpreted as an interest in the welfare of others, as players only aim to secure a high payoff for themselves.

In this framework, payoff maximisation is considered to be rational, but when such rational players then seize every opportunity to gain at the opponent's expense, they may counterintuitively both end up with low payoffs. A game that clearly exhibits this contradiction is the 'Traveler's Dilemma. Since its formulation in 1994 by the economist Kaushik Basu3, the game has become one of the most studied in the economics literature. Additionally, it has been discussed in theoretical biology in the context of evolutionary game theory.

In general, the dilemma relies on the individuals' incentive to undercut the opponent. To be more specific, players are motivated to claim a lower value than their opponent to reach a higher payoff at the opponent's expense. Such incentive leads players to a systematic mutual undercutting until the lowest possible payoff is reached, which is the unique Nash equilibrium. It seems paradoxical that players defined as rational in a game theoretical sense end up with such a poor outcome. Therefore, the question that naturally arises is how can this poor outcome be prevented and how cooperation can be achieved.

To address these questions, it can be helpful to better understand price wars, which consist in the mutual undercutting of prices to gain market share. In addition, it can provide information about human behaviour, because economic experiments have shown that individuals prefer to choose the cooperative high payoff action, instead of the Nash equilibrium4.

Our analysis focuses on showing that the Traveler’s Dilemma can be decomposed into a local and a global game. If the payoff optimisation is constrained to the local game, then players will inevitably end up in the Nash equilibrium. However, if players escape the local maximisation and optimise their payoff for the global game, they can reach the cooperative high payoff equilibrium.

1Max Planck Institute for Evolutionary Biology, Plön 24306, Germany. 2Max Planck Institute for Mathematics in the Sciences, Leipzig 04103, Germany. 3Santa Fe Institute for the Sciences of Complexity, Santa Fe, NM 87501, USA. *email: ramirez@mis.mpg.de
and $j$ is payoff and risk dominant\(^7,8\). In general, the equilibrium of the game, method to arrive to this equilibrium is called iterative elimination of dominated strategies\(^5\).

A Prisoner’s Dilemma. Traveler’s Dilemma consists of many embedded Prisoners’ Dilemmas. This means that at a local level the game is a Prisoner’s Dilemma (black boxes). This happens when the action space is reduced to any pair of actions $n$ and $n + s$, where $1 \leq s \leq R - 1$. While at the global level, we can observe a coordination game (gray boxes). This level is defined as any pair of actions $n$ and $n + s$, where $R \leq s \leq U - n$.

Here, we show that the cooperative equilibrium can be reached in a game like the Traveler’s Dilemma due to diversity, which we define as the presence of suboptimal strategies. The appearance of strategies far from those of the residents allows for the local maximisation process to be escaped, such that an optimisation at a global level takes place. Overall this can lead to cooperation because by considering “suboptimal strategies” that play against each other it is possible to reach higher payoffs, both collectively and individually.

**Game.** The Traveler’s Dilemma is a two-player game. Player $i$ has to choose a claim, $n_i$, from the action space, consisting of all integers on the interval $[L, U]$, where $0 \leq L < U$. The payoffs are determined as follows:
- If both players, $i$ and $j$, choose the same value ($n_i = n_j$), both get paid that value.
- There is a reward parameter $R > 1$, such that if $n_i < n_j$, then $i$ receives $n_i + R$ and $j$ gets $n_i - R$.

Thus, the payoff of player $i$ playing against player $j$ is

$$\pi_{ij} = \begin{cases} n_i & \text{if } n_i = n_j \\ n_i + R & \text{if } n_i < n_j \\ n_j - R & \text{if } n_i > n_j \end{cases} \quad (1)$$

Thus, a player is better off by choosing a slightly lower value than the opponent; when $j$ plays $n_j$, then it is best for $i$ to play $n_i - 1$. The iteration of this reasoning, which we will call the *stairway to hell*, leads to the only Nash equilibrium of the game, $[L, L]$, where both players choose the lowest possible claim. The classical game theory method to arrive to this equilibrium is called iterative elimination of dominated strategies\(^2\).

The game can be visualised through its payoff matrix (Fig. 1). For simplicity, we use the values from the original formulation: $L = 2, U = 100$ and $R = 2$. The payoff matrix shows that the Traveler’s Dilemma can be decomposed into a local and a global game. Let us begin with the local game. When the action space of the game is reduced to two adjacent actions $n$ and $n + 1$ (black boxes in Fig. 1), the Traveler’s Dilemma with $R = 2$ is equivalent to the Prisoner’s Dilemma\(^\text{a}\). In general, for any value of $R$, the Traveler’s Dilemma becomes a Prisoner’s Dilemma for any pair of actions $n$ and $n + s$, where $1 \leq s \leq R - 1$. For example, for $R = 4$ the pair of actions $n$ and $n + 1, n$ and $n + 2, n$ and $n + 3$ follow the same game structure as the Prisoner’s Dilemma. Therefore, the Traveler’s Dilemma consists of many embedded Prisoners’ Dilemmas. This means that at a local level the game is a Prisoner’s Dilemma.

If we now consider actions that are distant from each other in the action space, e.g. 2 and 100, we can observe a coordination game structure (gray boxes in Fig. 1), where $[100, 100]$ is payoff and risk dominant\(^\text{b}\). In general, any pair of actions $n$ and $n + s$, where $R \leq s \leq U - n$, construct a coordination game. As a result, the Traveler’s Dilemma becomes a coordination game at a global level, which has different equilibria than the local game.

**Paradox.** Social dilemmas appear paradoxical in the sense that self-interested competing players, when rationally playing the Nash equilibrium, end up with a payoff that clearly goes against their self-interest. But with the Traveler’s Dilemma, the paradox goes further, as suggested in its original formulation\(^7\). Classical game theory proposes $[L, L]$ as the Nash equilibrium of the game. However, it seems unlikely and implausible that, with $R$ being moderately low, say $R = 2$, for individuals to play the Nash equilibrium. This has been confirmed in economic experiments where individuals rather choose values close to the upper bound of the interval. Such experiments have also shown that the chosen value depends on the reward parameter ($R$), where an increasing...
Consequently, the aim of this paper is to seek and explore simple mechanisms through which the apparent non-rational cooperative behaviour can come about. We also examine the effect of the reward parameter on the game's outcome. Given that the Traveler’s Dilemma paradox emerges in the classical game theory framework, we analyse the game using evolutionary game theory tools. This dynamical approach allows us to explore adaptive behaviour outside of the stationary classical game theory framework. To be more precise, for this approach individuals dynamically adjust their actions according to their payoffs.

The key point of course is to understand how the system can converge to high claims. We show that this behaviour is possible because the Traveler’s Dilemma can be decomposed into a local and a global game. If the payoff maximisation is constrained to the local level, then the stairway to hell leads the system to the Nash equilibrium; given that locally the game is a Prisoner’s Dilemma. On the other hand, at a global level the game follows a coordination game structure, where the high claim actions are payoff dominant. Thus, for the system to reach a high claim equilibrium the maximisation process needs to jump from the local to the global level.

Our analysis led us to identify the mechanism of diversity as responsible for enabling this jump and preventing players from going down the stairway to hell. This mechanism works on the idea that to reach a high claim equilibrium, players have to benefit from playing a high claim. For a population setting, it means that players need to have the chance to encounter opponents also playing high. From a learning model point of view, it refers to the belief that the opponent will also play high, at least with a certain probability. If the belief is shared by both players, they should both play high and reach the cooperative equilibrium. Here, we explore these two types of models to unveil the mechanism leading to cooperation.

Population based models unveil diversity as the cooperative mechanism via the effect of mutations on the game’s outcome. This is shown for the replicator-mutator equation and the Wright–Fisher model. Similarly, a two-player learning model approach, more in line with human reasoning, shows that if players are free to adopt a higher payoff action from a diverse action set during their introspection process, they can reach the cooperative equilibrium. This result is obtained using introspection dynamics.

Finally, we explain how diversity is the underlying mechanism that enables the convergence to high claims in previously proposed models. To be more precise, we show that diversity is required because it allows for the maximisation process to jump from the local to the global level.

**Replicator–mutator equation**

**Formulation.** The replicator-mutator equation is a generalization of the replicator dynamics, which is a fundamental model to describe selection in a population. In replicator dynamics, the equation describes the evolution of an infinite population of \( n \) different types whose frequencies are \( x_1, \ldots, x_n \). To model selection, the reproduction rate of each type \( i \) is determined by its fitness \( f_i \), which is derived from the payoff matrix of interactions among individuals in the population\(^{14}\). In addition to the typical formulation of the replicator dynamics, the equation includes mutations via the mutation matrix, \( Q = (q_{ij}) \), where \( q_{ij} \) is the probability that type \( j \) mutates to type \( i \)\(^{14,15}\). Consequently, the mutation matrix \( Q \) is a row-stochasic matrix. Thus the replicator-mutator equation is described as

\[
\dot{x}_i = \sum_{j=1}^{n} x_j f_j(\bar{x}) q_{ji} x_i = -x_i f_i \quad i = 1, \ldots, n
\]

where \( \bar{x} = (x_1, x_2, \ldots, x_n) \) and \( f = \sum_{i=1}^{n} x_i f_i(\bar{x}) \) denotes the average population fitness.

To model mutations, we consider a uniform random probability of mutating to other types, i.e.

\[
q_{ij} = \frac{q}{n-1}, \quad i \neq j, \quad q_{ii} = 1 - q, \quad 1 \leq i, j \leq n
\]

The mutation parameter \( q \) can take values from 0 to \( \frac{n-1}{n} \). When \( q = 0 \), the dynamics corresponds to the replicator dynamics, such that no mutations occur. While for \( q = \frac{n-1}{n} \), type \( i \) is equally likely to change to any other type or to remain as it is\(^{14}\).

**Results.** The replicator dynamics limit (\( q = 0 \)) results are in agreement with classical game theory predictions. As expected, \( L \) is the claim that dominates the population regardless of the reward parameter value \( R \). Specifically, for this dynamics the stairway to hell takes place, where dominated strategies are systematically eliminated through extinction until only the Nash equilibrium claim remains in the population (Fig. 2a). This behaviour is related to the theorem proposed in Hofbauer and Sandholm\(^{15}\), where evolutionary game dynamics—like the replicator dynamics—successfully eliminate strictly dominated strategies, given that they fulfil a set of requirements.

When mutations are allowed (\( q \neq 0 \)), a different behaviour can be observed. The system reaches a coexistence equilibrium, where there is a claim with a higher frequency than the others. No extinction or dominance of actions takes place, due to the sustained presence of mutation events. Nonetheless, two types of this behaviour can be identified: for type 1 the system converges to a coexistence equilibrium where the frequency values of all claims are evenly distributed (Fig. 2b). For type 2, there is a single claim that nearly dominates the population given that its frequency is much larger than the rest (Fig. 2c).

To characterise how the system’s equilibrium changes according to the parameter values, the claim with highest frequency, the average claim and the average payoff were obtained for different different reward \( R \) and mutation strength \( (q) \) values. Figure 2d shows that the claim with highest frequency becomes larger for lower...
reward values ($c_{\text{highFreq}} \sim 1/R$). It also suggests that for reward values greater than $R_B \approx 50\sqrt{q}$, the claim with highest frequency is always 2. This result has been obtained by estimating the best fit for the boundary curve, due to the nature of the system of equations in question. As expressed by Eq. (2), the system consists of 99 coupled nonlinear differential equations. Consequently, a numerical approach is a natural method to find the solution of the system, and therefore a numerical fit is appropriate to estimate the boundary expression.

On the other hand, the results for the average claim of the population (Fig. 2e) imply that there is a boundary dividing the parameter space for which the system either exhibits type 1 or type 2. In particular, for $R \leq 20$, the boundary is $R_B \approx 50\sqrt{q}$, and for $R \geq 20$, the boundary is $R_B \approx 50q + 10$. If $R < R_B$, then type 1 is observed; while if $R > R_B$, then type 2 is observed. These boundaries were also obtained by estimating the best fit for the curves. The same results can be found via the average payoff of the population, defined as $\sum_{i=1}^{n_c} x_i f(x)$ (Fig. 2f). The reason for this similarity is the close relationship between the claims ($n_i$ and $n_j$) and the payoffs ($\pi_{ij}$), as shown in Eq. (1).

In general terms, the boundaries suggest that there is a critical mutation parameter value ($q$) for every reward value. This means that there is a minimum rate of mutations required for the maximisation process to jump from the local to the global level. As expected, for high reward values the system mainly converges to $L$, as the local level takes over the payoff matrix (see Fig. 1).

Overall, the results present the same outcome as economic experiments in terms of the reward parameter, where a low $R$ value is needed for high claims to be played most frequently. Additionally, they reveal that mutations are necessary for high claims to be advantageous in the population. This means that diversity should be maintained in the population for high claims to be favourable.

**Wright–Fisher model**

**Model.** The basic model for the evolution of populations including genetic drift, mutation and selection is the Wright–Fisher model, see e.g. [16–18]. It is therefore natural to also analyse the Traveler’s Dilemma within that framework. In particular, our aim is to further explore how the population can converge to high claims.

The Wright–Fisher model describes the evolution of a population with discrete, non-overlapping generations of fixed size $N$. From the current generation, the next generation is obtained by sampling with replacement, with
sampling weights proportional to the fitness of the individuals. For each sampled individual there is a probability \( \mu \) of mutation.

In our context, an individual is a player of the Traveler’s Dilemma with a certain value as her claim\(^20\), such that a mutation means that the value of the claim changes by \( \pm \delta \); unless this would result in a value smaller than \( L \) or larger than \( U \), in which case simply \( L \) or \( U \), respectively, is chosen.

The expected payoff of each individual is determined by playing the Traveler’s Dilemma with the other individuals present in the population, excluding self-interactions. The accumulated payoff \( \Pi \) then defines the fitness of an individual, that is, the sampling weight in the replacement scheme, via

\[
\text{fitness} = e^{\rho \Pi}
\]

where the selection intensity coefficient \( \rho \) modulates the magnitude of the impact of fitness differences in the game dynamics. The larger the selection intensity, the stronger the influence of fitness differences and the less important is the role of stochasticity on the dynamics.

The process of fitness dependent sampling with replacement and mutation probability \( \mu \) is repeated for \( t \) time steps. We fix the population size to \( N = 100 \) and consider \( \mu, \delta, \rho \) as variable parameters. We end the simulation after \( t = 1000 \) generations, because the dynamics converges well before that.

**Results.** For the Wright–Fisher model, a large mutation probability, \( \mu \), and a large maximal mutation size, \( \delta \), enable the convergence to high claims (see Fig. 3). The reason for this is that mutation introduces diversity in the population, larger mutation rates and sizes lead to greater diversity, and in particular, to the appearance of high claim individuals in the population. When players using these “suboptimal actions” play against each other, a higher payoff results, both collectively and individually. In general, a large mutation probability (\( \mu \)) and a large maximal mutation size (\( \delta \)) allow for convergence to high claims, given that these quantities are proportional to the diversity present in the population. Particularly, there is a boundary dividing the parameter space for which the system either converges to high claims or to \( L \), as shown by the dotted lines. For the strong selection regime, the system tends to converge to low claims because a high value of \( \rho \) reduces the effective population diversity by magnifying even small payoff differences. For very low \( \rho \) values, the level of stochasticity increases, given that the selection intensity parameter modulates the impact of fitness differences in the game dynamics. Finally, for high reward values the population converges to lower claims, as it is observed in economic experiments.
Additionally, the results show that for high reward values the population converges to lower claims, as it is also observed in economic experiments. In particular, the outcome of the simulations propose the following relationship between the parameters and the average claim: $c_{avg} \sim (\mu \delta)/(R \rho)$. Similarly, it suggests that there is a boundary dividing the parameter space for which the system either converges to high claims or to 1, defined by $\delta = \alpha \mu^{-b}$, where $\alpha$ and $b$ are constants such that $a \sim \delta \rho$ and $b \sim \rho$ (see dotted lines in Fig. 3). This boundary is a result of the maximisation process shifting from the local to the global level. As suggested by the previous expression, the eccentricity of the boundary is directly proportional to the selection intensity coefficient. The reason for this is that the selection intensity is inversely proportional to the population’s diversity, so for lower values of $\rho$ it is easier to reach a high claim equilibrium. However, when the maximal mutation size is low, the system cannot easily escape the local game and jump to the global game.

The boundary formula also suggests that when the reward value increases, the semi-major axis of the hyperbola that describes the boundary also increases, because for a greater reward value the region of the local game becomes larger (see Fig. 1). This means that greater values of $\delta$ and $\mu$ are needed to change from the local to the global maximisation process.

It should be noted that the average claim is equivalent to the average payoff of the population, given that a single claim dominates the population in the equilibrium.

On the whole, the Wright–Fisher model unveils that mere heterogeneity in the population is not sufficient to enable the convergence to high claims. This heterogeneity should be characterised by being diverse, i.e. there should be a presence of claims that are distant enough from each other in the action space, in order for the cooperative equilibrium to be reached. In fact, the model shows how the local maximisation process can be escaped thanks to diversity, introduced by mutations, in a population based model including genetic drift and selection.

### Introspection dynamics

**Model.** Introspection dynamics is a learning model used to describe adaptive behaviour via introvert reasoning. The basic idea is that a player, after considering her payoff in an encounter with an opponent, checks whether a randomly chosen alternative action—sampled from the uniform distribution on the $[L, U]$ interval—would have given her a better payoff against a given action chosen by her opponent. This process is called the introspection. A new action then is adopted with a probability depending on the payoff difference. Thus, when the payoff of the realised action is $\pi$, while the alternative action would have yielded $\tilde{\pi}$, the probability of adopting the latter action is computed using the Fermi function:

$$p = \frac{1}{1 + e^{-\beta(\tilde{\pi} - \pi)}}$$

with a selection intensity coefficient $\beta$. For weak selection ($\beta \to 0$), payoffs are not important in the learning process as any alternative action is randomly adopted. For strong selection ($\beta \to \infty$), players exclusively adopt the alternative action if their payoff matches at least the realised payoff. Therefore, $\beta$ determines the magnitude of the impact that payoffs have on the game dynamics. The parameter $\beta$ and the parameter $\rho$ of the Wright–Fisher model have different effects on the game dynamics, even though they share the same interpretation.

In this learning model, there are only two players that from time to time perform the introspection process, i.e. consider the adoption of a counterfactual action given that their opponent’s action is fixed. For the Traveler’s Dilemma, the model is used in the context of a one-shot game, where players adjust their behaviour using introspection.

While the proposed model may be stylised and coarse, the principle is not psychologically implausible, as humans can both evaluate putative alternatives and perform a coarse choice when the number of details in the decision process, here the number of possible claims, becomes too large.

The model thus assumes a rather simple cognitive mechanism, but as we shall see, it can lead to the emergence of cooperation. This is a novel approach and result in the context of the Traveler’s Dilemma, where noise and errors have been previously proposed as the cause for this behaviour.

**Results.** For the introspection dynamics, the system converges to large claims in the strong selection regime for low values of $R$ (see Fig. 4a). The dynamics, however, does not converge to a specific value, but instead oscillates in the high claim region. This is expected from the way the optimisation process operates. First, the global maximisation is performed, then the model shifts to a local maximisation. Given that at a local level the system does not have a maximum—due to the stairway to hell—the system exhibits oscillations in the large claim region. For the weak selection regime, as expected, the system becomes increasingly stochastic, until it comes to fluctuate across the whole action space (see Fig. 4b).

The stationary distribution of actions, derived using the main analytical result of Couto, Giorno and Hilbe, allows us to calculate the average claim and average payoff for different values of the reward parameter $R$ and the selection intensity coefficient $\beta$. As shown in Fig. 4c,d, strong selection and a low reward value are requirements for the convergence to high claims. This result is in agreement with economic experiments findings, where individuals choose high claims when the reward parameter is low. Such a decision is purely motivated by payoff maximisation, as in the case of strong selection, and not by mistakes or noise in the decision-making process.

To be more precise, for strong selection, the average claim follows the relation: $c_{avg} \sim 1/R$. In particular, $R_B \approx 9/\beta$ describes the boundary such that for $R > R_B$ the system always converges to 2. This boundary is obtained by estimating the best curve that fits the borderline. It should be noted that the complexity of the system is high; there are 99 possible actions per player. Thus, to compute the invariant distribution it is necessary to invert a 9801x9801 matrix, such calculation is done numerically, which makes a numerical fit an appropriate method to find the borderline. Overall, this boundary suggests that above a critical reward value the system converges...
to the lowest possible claim, because the maximisation process is constrained to the local level. This can be seen more clearly in Fig. 1, where it is explained that a larger $R$ increases the region of the local game. On the other hand, for the weak selection regime, the dynamics is characterised by stochasticity for all values of $R$. It should be noticed that the same general results hold for the average payoff because the payoff values are tightly related to the claim values, as explained by Eq. (1).

Once again it is found that for the convergence to large claims some diversity is needed, that is, the presence of a certain fraction of large claim players or a sufficiently high probability that agents select large claims. The random adoption of an action that outperforms the current one in the introspection dynamics generates and maintains such diversity.

Thus, the parameters $\rho$ and $\beta$ cause an overall different effect in the system’s dynamics. For the Wright–Fisher scenario, high values of $\rho$ prevent a global maximisation process by reducing the effective diversity of the population. On the other hand, $\beta$ does not have any effect on the diversity for the introspection dynamics, because the alternative random strategy is always chosen from a uniform distribution regardless of the magnitude of $\beta$. However, it should be noted that $\rho$ and $\beta$ share the same interpretation, as they are both parameters that modulate the impact of payoffs on the game dynamics. In particular, when $\rho$ or $\beta$ have low values, the level of stochasticity is large, which means that noise dominates the game dynamics, instead of the payoff differences.

Finally, we should note that, in contrast to our findings for the Traveler’s Dilemma, cooperation does not emerge in the Prisoner’s Dilemma when played using introspection dynamics. There, both players learn to defect at all times, that is, play Nash, in the limit of strong selection$^{22}$. Therefore, diversity is only effective in promoting cooperation when the structure of the game allows for it. The Traveler’s Dilemma is composed on a local and a global game, where jumping from the local to the global level maximisation is required to reach a cooperative equilibrium.
Comparison to other models

Previous models propose different methods to explain the convergence to high claims in the Traveler's Dilemma. From our perspective, however, they operate on the same underlying mechanism to allow for cooperation, namely diversity enabling a global maximisation to dominate over a local optimisation and thereby preventing the *stairway to hell*.

Two popular models are Goeree and Holt\(^26\), and Manapat et al.\(^27\). The first is a dynamic learning model with a logit rule, for which a player's decision probabilities are proportional to an exponential function of expected payoffs. In general, Goeree and Holt\(^26\) show that "noisy learning" allows for players to converge to high claims. This means that when players have the possibility to adopt non-Nash or "suboptimal" strategies, due to a "noisy learning process", the system can converge to high values. From our perspective, this is equivalent to having diversity in the game such that a global maximisation process occurs.

In comparison to the logit rule learning model, the introspection dynamics allows a better understanding of the paradox posed by the Traveler's Dilemma. The reason for this is that the introspection dynamics separates the notion of *rationality* from the mechanism of diversity. If it is assumed that a rational player is defined as a player that maximises its own payoff, then for strong selection ($\beta \rightarrow \infty$) players act rationally, as they always choose the claim with higher payoff in the introspection process. This allows for diversity to be introduced through the randomly chosen alternative action of the introspection step. Conversely, the logit rule introduces diversity via errors in the payoff maximisation.

In terms of evolutionary game theory models, in the stochastic approach formulated by Manapat et al.\(^27\) mutations lead to high claims. Overall, the results we obtained are in agreement with their results, as they also find it necessary to include mutations in their model to reach large values. Moreover, our and their approach exhibit the same behaviour in terms of the selection intensity coefficient. Nonetheless, in Manapat et al. diversity is not identified as the general underlying mechanism for the convergence to high claims. We have explained in the Wright–Fisher model section how diversity enables the jump towards cooperation in a population, when the Traveler's Dilemma is analysed using similar stochastic evolutionary dynamics.

Discussion

We have shown that diversity can enable cooperation to emerge in a game like the Traveler's Dilemma. This mechanism avoids the paradox of classical game theory where supposedly rational players inevitably end up in a Nash equilibrium, which neither maximises the individual nor the collective payoff. As already pointed out by Kaushik Basu\(^3\), and as confirmed by economic experiments\(^4\), such behaviour seems implausible for human players when faced with this game.

To analyse the dilemma, we proposed its decomposition into a local and a global game. Locally, the game structure is that of a Prisoner's Dilemma, while globally it is a coordination game. In consequence, for a cooperative high claim equilibrium to be reached, it is necessary to jump from the local to the global payoff maximisation. In order to explore this idea, we used a dynamical approach, including population based models and a learning model. The replicator-mutator equation and the Wright–Fisher model for the former, and introspection dynamics for the latter. Comparing the results allowed us to identify a general mechanism that may lead to the convergence to strategies far from the Nash equilibrium that leave the players much better off. Such a mechanism that favours cooperation is diversity. Here, diversity, understood as the presence of "suboptimal strategies", allows the jump from the local to the global level maximisation. To be more precise, there should be a sustained heterogeneity characterised by claims that are distant from each other in the action space. Overall this can lead to cooperation because by considering "suboptimal strategies" that play against each other it is possible to reach higher payoffs, both collectively and individually.

Additionally, we found that for all models the equilibrium claim is inversely proportional to the reward value ($c \sim 1/R$), this implies that $R$ delimits the competition between the local and global level. The stark borderline that appears in all models dividing the convergence to high claims and the convergence to the lowest possible claim is an evidence for this. Indeed, $R$ defines the region in the action space for which the Traveler's Dilemma is no longer a Prisoner's Dilemma, and becomes a coordination game. This result is in agreement with economic experiments, where it is shown that for increasing reward values the chosen claim decreases\(^5\).

This perspective also allowed us to identify diversity as the main mechanism of previous models used to analyse the Traveler's Dilemma\(^26,27\), although these models are very different.

Our approach may also help to better understand how humans play a game like the Traveler's Dilemma\(^4\), typically settling on high claims instead of going to the Nash equilibrium. While the introspection dynamics is rather stylised and coarse, it is not completely implausible as humans also tend to first make a coarse decision when faced with a large number of options. And in a game like the Traveler's Dilemma, a belief that opponents play "suboptimal" actions may become self-confirming when it leads to higher payoffs. From our perspective, it suffices that the probability to encounter a suboptimal high claim action is sufficiently high. This is ensured by diversity, as previously explained.

Overall, the mechanism of diversity proposed here benefits players both at a collective and at an individual level. Therefore, exploring other games that share the same characteristic of having conflicting interests between a local and a global payoff maximisation might be helpful to understand and induce cooperation in scenarios such as price wars. A situation where undercutting the opponent is tempting but reaching a cooperative equilibrium would be beneficial for everybody.

Data availability

The code for the simulations and numerical solutions is available at [https://github.com/MA-Ramirez/Travelers_Dilemma_Code.git](https://github.com/MA-Ramirez/Travelers_Dilemma_Code.git).
Received: 25 October 2022; Accepted: 20 January 2023
Published online: 25 January 2023

References
1. Fudenberg, D. & Tirole, J. Game Theory 6th edn. (MIT Press, Cambridge, 1998).
2. Broom, M., Pattni, K. & Rychtář, J. Generalized social dilemmas: The evolution of cooperation in populations with variable group size. Bull. Math. Biol. 81, 1 (2018).
3. Basu, K. The traveler's dilemma: Paradoxes of rationality in game theory. AEA Pap. Proc. 84, 391–395 (1994).
4. Capra, C. M., Goeree, J. K., Gomez, R. & Holt, C. A. Anomalous behavior in a traveler’s dilemma. Am. Econ. Rev. 89, 678–690 (1999).
5. Hofbauer, J. & Sigmund, K. Evolutionary Games and Population Dynamics (Cambridge University Press, Cambridge, UK, 1998).
6. Rapoport, A. & Chammah, A. M. Prisoner's Dilemma (University of Michigan Press, Ann Arbor, 1965).
7. Fudenberg, D. & Levine, D. The Theory of Learning in Games (MIT Press, Cambridge, MA, 1998).
8. Bayer, P. et al. Coordination games in cancer. PLoS One 17, 1 (2022).
9. Maynard Smith, J. & Price, G. R. The logic of animal conflict. Nature 246, 15–18 (1973).
10. Weibull, J. W. Evolutionary Game Theory (MIT Press, Cambridge, 1995).
11. Roca, C. P., Cuesta, J. A. & Sanchez, A. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. Phys. Life Rev. 6, 208–249 (2009).
12. Broom, M. & Rychtář, J. Game-Theoretical Models in Biology (Chapman and Hall/CRC, 2013).
13. Bauer, J., Broom, M. & Alonso, E. The stabilization of equilibria in evolutionary game dynamics through mutation: mutation limits in evolutionary games. Proc. R. Soc. A 475, 1 (2019).
14. Duong, M. & Han, T. A. On equilibrium properties of the replicator-mutator equation in deterministic and random games. Dyn. Games Appl. 10, 641–663 (2020).
15. Hofbauer, J. & Sandholm, W. Survival of dominated strategies under evolutionary dynamics. Theor. Econ. 6, 341–377 (2011).
16. Imhof, L. A. & Nowak, M. A. Evolutionary game dynamics in a Wright-Fisher process. J. Math. Biol. 52, 667–681 (2006).
17. Otto, S. P. & Day, T. A Biologist’s Guide to Mathematical Modeling in Ecology and Evolution (Princeton University Press, Princeton, New Jersey, 2007).
18. Etheridge, A. Some mathematical models from population genetics Lecture notes in mathematics (Springer, Berlin Heidelberg, 2011).
19. Hofrichter, J., Jost, J. & Tran, T. Information geometry and population genetics: the mathematical structure of the Wright-Fisher model Understanding complex systems (Springer, Cham, 2017).
20. Datseris, G., Vahdati, A. R., DuBois, T. C. & Agents.IL A Performant and Feature-Full Agent-Based Modeling Software of Minimal Code Complexity (Sage Publications, 2022).
21. Hauser, O. P., Hilbe, C., Chatterjee, K. & Nowak, M. Social dilemmas among unequals. Nature 572, 524–527 (2019).
22. Couto, M. C., Giarmo, S. & Hilbe, C. Introspection dynamics: A simple model of counterfactual learning in asymmetric games. New J. Phys. 24, 1 (2022).
23. Blume, L. E. The statistical mechanics of strategic interaction. Games Econ. Behav. 5, 387–424 (1993).
24. Szabó, G. & Tóke, C. Evolutionary Prisoner’s Dilemma game on a square lattice. Phys. Rev. E 58, 69–73 (1998).
25. Hauert, C. & Szabó, G. Game theory and physics. Am. J. Phys. 73, 405–414 (2005).
26. Goeree, J. K. & Holt, C. A. Stochastic game theory: For playing games, not just for doing theory. Proc. Natl. Acad. Sci. USA 96, 10564–10567 (1999).
27. Manapat, M., Rand, D., Pawlowitsch, C. & Nowak, M. Stochastic evolutionary dynamics resolve the traveler’s dilemma. J. Theor. Biol. 303, 119–127 (2012).

Acknowledgements
We would like to thank Christian Hilbe and Marta Couto for valuable feedback and discussions.

Author contributions
M.A.R. conceived the study, J.J. and M.S. contributed. M.A.R. and M.S. performed the computer simulations. M.A.R. computed the numerical solutions. All authors analysed and discussed the results. M.A.R. and J.J. wrote the manuscript. All authors reviewed the manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to M.A.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2023