Technical Note

The Tragedy of the Commons as a Prisoner's Dilemma. Its Relevance for Sustainability Games

Alessio Carrozzo Magli [1,*]**, Pompeo Della Posta** [2] **and Piero Manfredi** [2]

1   Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy
2   Department of Economics and Management, University of Pisa, Via Ridolfi 10, 56124 Pisa, Italy; pompeo.della.posta@unipi.it (P.D.P.); piero.manfredi@unipi.it (P.M.)
*   Correspondence: alessio.carrozzo2@unibo.it

Abstract: In the current battle for sustainability and climate, understanding the nature of sustainability games is of paramount importance, especially to inform appropriate policy actions to contrast the harmful effects of global climate change. Relatedly, there is no consensus in the literature on the proper game-theoretic representation of the so-called Tragedy of the Commons. A number of contributions have questioned the prisoner's dilemma as an appropriate framework. In this work, we provide a representation that reconciles these two positions, confirming the ultimate nature of the Tragedy as a prisoner's dilemma, rather than a coordination issue, and discuss the ensuing implications for sustainability policy interventions.

Keywords: sustainability game; prisoner's dilemma; coordination game; Tragedy of the Commons; green investment; Nash equilibrium

1. Introduction

The planetary challenge of global climate change and the planned contrasting actions such as, e.g., in the case of Europe, the European Green Deal [1], as well as the momentum to the ecological transition that should be promoted by the post-pandemic recovery plans, will require an unprecedented supporting effort by science, at any level. In relation to this, a fundamental step on the theoretical side would be to achieve a full understanding of the nature of what is termed here as a "sustainability game". By "sustainability game", we mean any form of game, that is, any strategic interactions—involving any type of economic, social and political agents (from consumers and producers, to intermediate societal bodies up to institutions and governments)—specifically dealing with the environment *latu sensu* whose outcome could have a relevant impact on the Earth's fundamental resources that will be available to current and future generations. With special reference to climate change, Heitzig et al. [2] spoke instead of a "mitigation game". This definition obviously includes the current battle against global climate change as its ultimate and pervasive endpoint. A number of contributions have identified the prisoner's dilemma as the most suited game-theoretic representation of sustainability games (e.g. Ostrom [3], Cooper and Barrett [4], Heugues [5], Dutta and Radner [6], Heitzig et al. [2]) other works have focused on coalition formation as a strategy to cope with the resulting free-riding problem (Hannam et al. [7], Nordhaus [8], Finus [9], Carraro and Siniscalco [10] and Wood [11] and a few have underlined the role of politics and institutions in strategic interactions (North [12]). Surely, an indisputable merit of the prisoner's dilemma is that of lucidly highlighting the tension between individual selfish behavior and collective interest, where everyone, pursuing the former, achieves an overall lower welfare than cooperation would instead ensure. The prisoner's dilemma tends to arise conceptually in whatever type of strategic interactions. For example, consider two producers that have to choose between a clean ("sustainable") and a dirty process. The latter (the former) is cheaper (more expensive) but has an (no) impact on the environment. From an individual standpoint, it is convenient to adopt

the dirty process to minimize costs (thereby maximizing profits) in spite of forcing the community to share an environmental cost (i.e., a negative externality). Since this is true for both (all) players, all players will play "dirty". However, the doubled environmental damage at the community level will compensate the individual increase in profits, such that eventually, had both adopted the clean process, individual welfare would be higher for all players, as well as for the community as a whole that preserved its environment. It is easy to realize that such considerations apply, mutatis mutandis, to any other instance of strategic interactions between economic or political agents (consumers, governments, intermediate bodies, etc.) facing issues of profit maximization (or cost minimization) in an interactive setting. Notably, all the previous ones are instances of the so-called free rider problem (Stiglitz [13]) that arises when considering public and common goods.

In general, however, the consensus on this is far from being unanimous as, for example, recent works have rather pinpointed the coordination nature of climate negotiations (DeCanio and Freimstand [14], Barrett [15,16]) as well as of more general economic games dealing with the environment (Mielke and Steudle [17]).

A much-related concept, stemming directly from the (un)sustainability of natural resource exploitation beyond the Earth's natural regenerative ability and carrying capacity, is the classic problem of the Tragedy of the Commons, first introduced by Lloyd [18] and later popularized by Hardin [19]. By their very definition, common goods, such as free natural resources, satisfy the well-known properties of non-excludability and rivalry. However, those who have access to them cannot resist the temptation to over-exploit them at the expense of other people and the community as a whole. In his seminal work, Hardin made the example of a wide pasture feeding the cows belonging to a large community. Under these circumstances, each agent has a private incentive to add a cow to increase individual profit because he perceives that this additional cow will have a small (negative) impact on the amount of pasture available to the community as a whole, what we will term, just for organizing the discussion, the first Hardin postulate. Therefore, each agent will add cows and the process will unavoidably continue until the common will be completely exhausted, what we will term later on as the second Hardin postulate.

As with the prisoner's dilemma, the Tragedy of the Commons also effectively demonstrates the tension between individual selfish behavior and the collective interest.

The relationship between the Tragedy of the Commons and the prisoner's dilemma has attracted the attention of scholars from different disciplines. Some early works (Dawes [20] and references therein, Olson [21] and Ostrom [3]) simply postulated that the Tragedy of the Commons is a prisoner's dilemma with a large number of players. Based on a number of celebrated field experiments, Rapoport [22] concluded that cooperation in a four-player Tragedy of the Commons is far more difficult to achieve than in a two-player prisoner's dilemma. This finding is not at odds with the interpretation of the Tragedy as a prisoner's dilemma, given that the increase in the number of players extends the relevance of the prisoner's dilemma and worsens its impact (Ostrom [3]). An intuitive explanation of this circumstance lies in two main arguments: (i) the fact that the larger the number of agents, the larger the communication issues will be, thereby making coordination harder and harder, and (ii) the fact that the larger the number of agents, the lower both the perceived and the actual environmental impact of each single player will be. In relation to point (ii), think, for instance, of an oligopoly of dirty producers in a fixed size economy. In this case, the higher the number of firms, the lower individual production and therefore the lower the individual impact on the environment, in spite of an increasing aggregate production. However, the tension between the prisoner's dilemma and coordination reappears here as well. For example, Barrett [16], and Decanio and Fremstad [14] implicitly recalled the Tragedy of the Commons, when they argued that the prisoner's dilemma, typical of climate negotiations and sustainability games, could evolve into a coordination game when players have perfect and symmetric information about the threshold of environmental damage that triggers a climate catastrophe. From a quite different perspective, Romagny et al. [23] claimed that the two-player Tragedy of the Commons loses its nature of a pris-

oner's dilemma when the set of strategies is continuous rather than binary. Through a more articulated discussion in this journal, Diekert [24] suggested a number of factors undermining the prisoner's dilemma's interpretation of the Tragedy. First, he pinpointed that when the number of players is large, the influence of each single individual becomes essentially irrelevant, meaning that the problem loses its strategic nature. More importantly, if one attempts at maintaining the strategic nature, the problem would rather point towards a coordination issue. Last, he suggested that in order to correctly capture Hardin's reasoning from an ecological standpoint, it is necessary to explicitly include the dynamics of the involved natural resource, thereby ending in a complicated problem and losing the possibility to have a simple game-theoretic description.

In this article, we aim to reconcile these different positions by showing that, unlike Diekert's approach, the interpretation of the Tragedy of the Commons as a prisoner's dilemma generally holds under fairly natural hypotheses, namely, when each player's impact is negligible, and behaviors that are disruptive towards the environment are more profitable from an individual standpoint. In particular, we provide two proofs: an elementary one, non-technical, in the main text, and a more rigorous one in Appendix A.

2. Modeling the Tragedy

2.1. Diekerts's Representation of the Tragedy of the Commons

We strictly follow Diekert's formulation, as reported in his Figure 3 (Diekert [24], p. 1779), which aimed to closely follow Hardin's narrative (Table 1). There are many identical players feeding their cows on a common finite pasture. If all players cooperate (C, for brevity), having only one cow each, everybody receives a unit payoff (clearly, one could use any positive constant). However, any player (call it "Player 1" just to fix ideas) can defect (D), by adding another cow to the pasture, and share the ensuing cost with all the others, thereby receiving an almost double payoff (2-ε), while all the others would receive a payoff of (1-ε), where ε is the per capita cost on the assumption that the cost of environment over-exploitation is uniformly shared by all players. Since all players face the same problem, when all defect, the outcome will be a zero payoff for everyone. (Note that the symmetry in the over-exploitation of resources can determine zero payoffs when considering: (i) a large population (possibly close to the environmental carrying capacity), meaning that each agent has a seemingly negligible impact on the outcome, as in the classical representation of the Tragedy; (ii) a few "large-scale" agents (in the proposed interpretation, at least three, but a more realistic generalization is able to also encompass the two players of the original dilemma). In particular, if $N = 2$, then zero payoffs would be obtained with $\varepsilon = 1$. By slightly increasing the private benefit from adding a cow, a standard prisoner's dilemma emerges. The points (i) and (ii) above allow further appreciating the relationship between the Tragedy and the prisoner's dilemma).

Diekert's main conclusion, resulting from Table 1, is that the strategy where everybody defects does not represent a Nash equilibrium, which rules out the nature of a prisoner's dilemma. Rather (still from Table 1), players are facing a problem of coordination. Indeed, Table 1 actually represents an anti-coordination (actually, a "chicken") game, since the equilibria in the pure strategy are the configurations in which Player 1 defects and all other players cooperate, and vice versa. The most striking consequence of Diekert's setup is that the Nash equilibrium (C,D) implies that the cooperation of a single player is sufficient to fully avoid the Tragedy of the Commons.

Table 1. Payoff matrix of Hardin's game-theoretic representation of the Tragedy of the Commons according to Diekert's Figure 3 (Diekert [24], p. 1779).

Player 1	All Other Players	
	Cooperate	**Defect**
Cooperate	$1, 1$	$1 - \varepsilon,\ 2 - \varepsilon$
Defect	$2 - \varepsilon, 1 - \varepsilon$	$0, 0$

2.2. Shortcomings of Diekerts's Representation and an Alternative Formulation Removing Them

The main drawback with Diekerts's argument is that his representation of the game is highly asymmetric: if Player 1 defects, the cost he imposes on each of the other players is the same imposed by all other players (whose individual impact would instead amount to $\frac{\varepsilon}{N-1}$). His hypothesis therefore implies a contradiction: either there are only two symmetric players, or—if we consider a multi-player setting—Player 1 has a far more than negligible impact, contradicting Hardin's (and Diekert's) first postulate.

To remove this contradiction, let us rebuild the relevant payoff matrix (Table 2), first, by restoring symmetry, assuming that all agents have the same impact (we keep the notation ε) on the environment, second, by imposing Hardin's second postulate, namely, that if all players defect by adding a cow each (this means, following Hardin, that each player expects an abrupt unit increase in their payoffs), the payoff becomes zero for everyone. Therefore, in this case, since each agent's payoff fulfils the equality $2 - \varepsilon N = 0$, it holds that $\varepsilon = \frac{2}{N}$.

In what follows, we set these simple ideas within Diekert's framework, which is appealing in view of its simplicity, to provide a very direct and straightforward proof of the prisoner's dilemma of the Tragedy.

Note, first, that cells (C,C), (D,C) and (D,D) in Table 2 are equivalent to Diekert's formulation. As for cell (C,D),when only Player 1 cooperates and all others defect, it follows that there are $N - 1$ agents, each one imposing the cost ε. The resulting payoffs will then be $1 - (N-1)\varepsilon = 1 - \frac{2(N-1)}{N} = -\frac{N-2}{N}$ (which approaches -1 for large N) for Player 1, and $2 - \varepsilon(N-1) = 2 - \frac{2(N-1)}{N} = \frac{2}{N}$ (which approaches 0 for large N) for all other players.

Table 2. Payoff matrix of the "corrected" game-theoretic representation of the Tragedy of the Commons when each player imposes a cost ε on the others.

	All Other Players	
Player 1	**Cooperate**	**Defect**
Cooperate	$1, 1$	$-\frac{N-2}{N}, \frac{2}{N}$
Defect	$2 - \frac{2}{N}, 1 - \frac{2}{N}$	$0, 0$

At a first glance, Table 2 seems to suggest that the pair (Defect, Cooperate) is a Nash equilibrium. However, this conclusion is wrong. The mistake stems from considering "all other players" as a unique agent or, alternatively, as a multitude of agents playing cooperatively within their group and non-cooperatively against Player 1.

To avoid this shortcoming, let us proceed in the most plain manner by explicitly disentangling the role of "other players" in the payoff matrix. (The representation in Table 1 implicitly assumes a zero lower bound on payoffs, occurring when the pasture is completely exhausted. This is at odds with Tables 2 and 3, where the "sucker's payoff" (i.e., the payoff for the player choosing to cooperate while the others defect) becomes negative and approaches -1. There are several ways to address this problem. First of all, setting a lower bound to 0 for the sucker's payoff only transforms the game into a weak prisoner's dilemma, but rational players would still play defect. Results would be unchanged. Second, the Tragedy of the Commons does not necessarily imply the complete depletion of the natural resource. It is sufficient that the configuration where everyone defects is Pareto-dominated by the configuration where everyone cooperates to conclude that the resource is over-exploited. Last, when all players but one defect, the common good is already doomed to exhaustion. In this context, it might well be that players will try to exploit, as much as possible, what is left before the complete collapse, at least achieving a temporary private benefit.) This implies considering a spectrum of possibilities depending on the number of agents belonging to the "other players" group that are either cooperating

or defecting (Table 3). In this correct formulation, Player 1 has the incentive to defect. Since all players face the same problem, each of them will defect.

Table 3. The explicit *N*-player game-theoretic representation of the Tragedy of the Commons based on splitting the $(N-1)$ "other players" into subgroups of cooperators and defectors. Different colors and styles are used to denote the payoffs of the various agents: "Player 1" (black), "other players" that do cooperate (red), "other players" that defect (bolded blue).

		All Other Players (*N*−1)			
		N−1 Cooperate (No One Defects)	*N*−2 Cooperate, 1 Defects	. . .	All Defect
Player 1	Cooperate	$1, 1, \ldots 1$	$1 - \frac{2}{N}, 1 - \frac{2}{N} \ldots, 1 - \frac{2}{N},$ $2 - \frac{2}{N}$	. . .	$-\frac{N-2}{N},$ $\frac{2}{N}, \ldots, \frac{2}{N}$
	Defect	$2 - \frac{2}{N},$ $1 - \frac{2}{N}, \ldots, 1 - \frac{2}{N}$	$2 - \frac{4}{N},$ $1 - \frac{4}{N}, \ldots, 1 - \frac{4}{N},$ $2 - \frac{4}{N}$	. . .	$0,$ $0, \ldots, 0$

In particular, it is easy to check that the action "Defect" strictly dominates (i.e., it always yields a higher payoff) the action "Cooperate" for Player 1, regardless of the actions played by the other players, and the same holds for all the other $(N-1)$ players. This result is a consequence of the assumption that the private benefit from adding a cow (1) is higher than the corresponding (individual) cost $(\frac{2}{N})$. The unique Nash equilibrium (and unavoidable outcome) of the game occurs when each player defects, in turn implying that Table 3 (and consequently Table 2 when properly interpreted) correctly represents an *N*-player prisoner's dilemma.

In Appendix A, interested readers can find a running example based on the case $N = 3$ as well as rigorous proof supplementing the elementary approach proposed here.

3. Discussion

The very recent, extraordinary discovery that micro-plastics can invade even human placentas (Ragusa et al. [25]), and the conjecture reported therein that this might occur via the food chain due to the dramatic extent of oceans' micro-plastic pollution (Landrigan et al. [26]), is just Earth's last, but utmost, signal that humanity is passing the point of no return. The disruption of many oceanic environments, and the dramatic impact of the increasing temperature on poor economies (Dell et al. [27]) are global instances of the Tragedy of the Commons, preconized by many (starting from Lloyd [18] but formalized by Hardin [19]).

Motivated by the need to improve our theoretical understanding of the nature of sustainability games for policy purposes, in this note, we departed from a previous synthesis that appeared in this journal (Diekert [24]), to reconsider the Tragedy of the Commons from a game-theoretic perspective. The direct implications of our analysis are threefold. First, it shows that the Tragedy of the Commons represents a true strategic interaction, contrary to Diekert's claim that the strategic nature is lost (ending in a coordination problem instead) due to the negligible impact of individual agents. Second, the Tragedy of the Commons robustly emerges as a true *N*-player prisoner's dilemma, even for large (but finite) *N*. Last, it suggests that a meaningful game-theoretic analysis of the Tragedy can be provided regardless of the dynamics of the natural resources or eco-systems specifically involved, as instead suggested in Diekert [24], therefore providing a greater level of generality.

Notably, under highly specific circumstances, the Tragedy of the Commons can also be given a structure of a hawk–dove game. Nevertheless, for this to be the case, the impact of ε on players' payoff must have a discontinuity, similar to what was hypothesized in [16]. This corresponds to the idea that a few defecting players should have a small impact on payoffs up to a certain threshold, beyond which any further, even minimal, defections might cause the sudden collapse of the natural resource or of the ecosystem as a whole. We believe this was an implicit feature of the Diekert setup.

Additionally, as resulting from the present work, in general, the anti-coordination nature (i.e., chicken game) does not seem to be representative of the one-shot interactions typical of standard sustainability games, which instead appear as well-defined prisoner's dilemmas. Coordination might, at most, represent an emerging dynamic feature, i.e., one emerging over time after many shots of the underlying prisoner's dilemma results in defections. In this case, the extent of the cumulative environmental degradation would be so dramatic that even a few further defections will be sufficient to trigger a catastrophic collapse, in line with [16].

The possibility that coordination emerges at an earlier stage compared to this highly degraded scenario requires non-myopic players capable to sustain the cooperative outcome over a repeated game, for instance, by adopting a grim trigger strategy. Sadly, the remarkable failures of climate agreements starting from Rio de Janeiro 1992 up to COP25 in Madrid are an even-too-clear sign that the required level of agents' (at all levels) far-sightedness is still dramatically insufficient (Walker et al. [28], Bang et al. [29], Clémençon [30]).

To sum up, we believe that the present clarification on the nature of sustainability games, particularly the Tragedy of the Commons, as a prisoner's dilemma, rather than a coordination issue—as claimed in recent research—is of paramount importance to inform the policy actions against global climate change. The need to always identify the ultimate nature of strategic interactions should be considered in all situations where sustainability issues arise. Indeed, if the nature of the game is purely coordinative, it is sufficient to enact adequate communication to achieve cooperation. In contrast, if the nature of the game is that of a prisoner's dilemma, communication alone would plainly be insufficient, and a third party, e.g., a government intervention (or an infinitely repeated game with patient enough agents) would be required.

4. Conclusions

The potentially ubiquitous nature of the prisoner's dilemma should never be forgotten while planning any form of intervention aimed to mitigate the impact on climate. Indeed, even though the ability to enact coordination at many different levels and scales will be a key resource for a successful battle against climate change, undue emphasis on a merely coordinative nature of sustainability games might be deleterious in view of the implicit optimistic message that it conveys. Instead, the correct acknowledgement of the main sustainability games—including the Tragedy of the Commons—as a full prisoner's dilemma, would unambiguously call for rapid governmental measures aimed to correct the resulting negative externality in the short time scales that current projections on climate change are clearly indicating.

Appendix A

Appendix A.1. A Running Example with N = 3 Agents

For an easy catch of our main result, here, we rearrange Table 3 of the main text in a simplified setting with $N = 3$ agents and therefore $\varepsilon = \frac{2}{3}$. The reader immediately sees that "alter" is, in this case, represented by the other two players, and it expands into four sub-cases depending on whether Player 2 (3) cooperates or defects. Note that for obvious

reasons of symmetry, one could discard one of the central columns, but we report all of them for full clarity.

Table A1. Explicit 3-player game-theoretic representation of the Tragedy of the Commons. Different colors and styles are used to denote the payoffs of the various agents: "Player 1" (black), "Player 2" (red), "Player 3" (bolded blue).

		All Other 2 Players			
		Player 2 Cooperates Player 3 Cooperates	Player 2 Defects, Player 3 Cooperates	Player 2 Cooperates Player 3 Defects	Player 2 Defects Player 3 Defects
Player 1	Cooperate	$1, 1, 1$	$\frac{1}{3}, \frac{4}{3}, \frac{1}{3}$	$\frac{1}{3}, \frac{1}{3}, \frac{4}{3}$	$-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}$
	Defect	$\frac{4}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}$	$\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}$	$0, 0, 0$

One can easily see that the strategy "Defect" strictly dominates the strategy "Cooperate" for Player 1 (and therefore for all players), meaning that, eventually, all players will defect. Indeed, focusing, e.g., on Player 1, the payoffs reported in the "Defect" line are always strictly greater than those reported in the "Cooperate" line. In doing so, we carried out a simple application of the strict dominance criterion used for elementary games.

Appendix A.2. A Rigorous Proof of the Main Result

Here, we provide a general proof of our main statement that the Tragedy of the Commons is always a prisoner's dilemma, provided that it holds Hardin's first postulate that each individual impact on the common good is small. The proof extends to the present case a general game-theoretic argument used in the literature (see e.g., Bauch and Earn 2004, [31]).

Consider, for simplicity, an individual (Player 1 of the main text) who faces a binary choice, namely, to either add one cow with probability P, or two cows, with the complement probability $1 - P$ (in this way, the expected number of cows is a continuous function between the two arbitrary boundaries, also addressing the point made by Romagny et al., 1997). Conventionally, we say probability P represents the individual's strategy. Assume that a single cow ensures a profit of u, while two cows ensure a higher profit $v > u$. Let π_n represent the probability that the pasture is over-exploited up to exhaustion, leaving all players with a payoff of 0, taken as a function of the proportion n of players adding only one cow. We make the following assumptions on π_n:

(i) $\pi_n \langle \pi_0 \leq 1 \,\forall n \rangle 0;$

(ii) $\frac{\partial \pi_n}{\partial n} < 0 \,\forall n < n^{crit}$;

(iii) $\frac{\partial \pi_n}{\partial n} = 0 \,\forall n > n^{crit}$;

(iv) $\pi_{n^{crit}} = 0.$

Assumption (i) is straightforward: when all N players add two cows, the probability that the pasture is exhausted is maximized and may also be the sure event.

Assumption (ii) means that the probability that the pasture is exhausted is a decreasing function of the proportion of "clean" players (i.e., those adding only one cow). Assumptions (iii) and (iv) mean that when many players are clean, the probability of exhaustion becomes 0. Our results do not depend on whether π_0 is exactly equal to one rather than strictly lower than 1, but thus far, the discussion is based on the fact that $\pi_0 = 1$, in order to adhere to Hardin's second postulate. For the same reason, we can also state that $n^{crit} < 1$.

The expected payoff of an individual playing strategy P, i.e., adding one cow with probability P, also obviously depends on the strategies played by other individuals as well. Therefore, one should look at

$$E(P, \pi_n) = u(1 - \pi_n)P + v\left(1 - \pi_{n-N^{-1}}\right)(1 - P) \qquad (A1)$$

where the notation $n - N^{-1}$ is used to mean that, if a player switches from one cow to two cows, the fraction of players adding one cow falls by $\frac{1}{N} = N^{-1}$.

The last expression is equivalent to the payoff matrix in Table 3. When Player 1 adds one cow only, their payoff u (equal to 1 in Table 3) also depends on the interaction with the other players: the higher the number of players adding two cows (i.e., the lower n), the lower their profit will be (because of the assumption $\frac{\partial \pi_n}{\partial n} < 0$), whereas, in Table 3, the individual payoff (1) is reduced by $\frac{2}{N}$ each time a player adds a cow (meaning that it still decreases as n decreases). When Player 2 adds two cows, their payoff v (equal to 2 in Table 3) again decreases in the number of players adding two cows (including Player 1 themself, since π_n falls to $\pi_{n-N^{-1}}$, which is equivalent to the cost $\frac{2}{N}$ that Player 1 caused to all N players when defecting). In a few words, the cost $\frac{2}{N}$ stemming from a further cow reduces the expected payoff in Table 3, exactly as the reduction of π_n to $\pi_{n-N^{-1}}$ does. Now, let us define, by $z = \frac{u}{v} < 1$, the relative profit resulting from adding one cow, meaning that we can rewrite the previous expression as

$$E(P, \pi_n) = (1 - \pi_n)Pz + (1 - \pi_{n-N^{-1}})(1 - P) \tag{A2}$$

We are now going to prove that there exists a unique Nash equilibrium (which is also convergently stable) irrespective of the functional form of π_n and of the actual values (u, v), provided only that $v > u$, and π_n is flat enough (i.e., each agent has a small impact on the overall probability of exhaustion).

Suppose that a fraction x of the population adds one cow with probability P and the rest adds one cow with probability Q, meaning that the proportion of adding one cow is $n = [xP + (1 - x)Q]$.

The (expected) payoff for an individual playing P is therefore

$$E_P(P, Q, x) = \left(1 - \pi_{[xP+(1-x)Q]}\right)Pz + \left(1 - \pi_{[xP+(1-x)Q]-N^{-1}}\right)(1 - P) \tag{A3}$$

Meanwhile, the (expected) payoff from an individual playing Q is

$$E_Q(P, Q, x) = \left(1 - \pi_{[xP+(1-x)Q]}\right)Qz + \left(1 - \pi_{[xP+(1-x)Q]-N^{-1}}\right)(1 - Q) \tag{A4}$$

The expected payoff gain for an individual playing P instead of Q is therefore given by the difference

$$\Delta E = E_P(P, Q, x) - E_Q(P, Q, x) = (P - Q)\left[\left(1 - \pi_{[xP+(1-x)Q]}\right)z - \left(1 - \pi_{[xP+(1-x)Q]-N^{-1}}\right)\right] \tag{A5}$$

Now, according to Hardin's first postulate (and, of course, to Diekert) that individual activities are negligible for the overall pasture, $\pi_{N[xP+(1-x)Q]} \approx \pi_{N[xP+(1-x)Q]-N^{-1}}$, the previous expression can be approximated by

$$\Delta E \approx (P - Q)\left(1 - \pi_{[xP+(1-x)Q]}\right)(z - 1) \tag{A6}$$

For any given z, there is a unique strategy $P = P^*$ such that $\Delta E > 0$ for any other strategy $Q \neq P^*$ and for any proportion x, and this is equal to $P^* = 0$, meaning that $P^* = 0$ is the unique Nash equilibrium. Indeed, since $z < 1$, it follows that $\left(1 - \pi_{[xP+(1-x)Q]}\right)(z - 1)$ is always negative. The only way to make ΔE always positive is to have Q always lower than P^* for any $Q \neq P^*$, and this is the case only if $P^* = 0$. However, $P^* = 0$ means that all players add two cows, meaning that they eventually obtain a payoff of 0. Had anyone added one cow only, each would have obtained a payoff of $u > 0$. This is exactly the definition of an N-player prisoner's dilemma.

As a final remark, notice that the coordinative nature would appear only if the above approximation $\pi_{[xP+(1-x)Q]} \approx \pi_{[xP+(1-x)Q]-N^{-1}}$ were not correct because in that case, we could find a mixed strategy equilibrium where the proportion of clean players is positive, but this simply confirms the point we already stressed when criticizing Diekert's

approach: the coordinative nature emerges only when one assumes that each player has a non-negligible impact, which is in contrast to Hardin's postulates of the Tragedy of the Commons.

References

1. The European Green Deal a European Green Deal, Striving to Be the First Climate-Neutral Continent. Available online: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en (accessed on 28 November 2020).
2. Heitzig, J.; Lessmann, K.; Zou, Y. Self-enforcing strategies to deter free-riding in the climate change mitigation game and other repeated public good games. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15739–15744. [CrossRef] [PubMed]
3. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*; Cambridge University Press: Cambridge, UK, 1990.
4. Cooper, R.N.; Barrett, S. Environment and Statecraft: The Strategy of Environmental Treaty-Making. *Foreign Aff.* **2003**, *82*, 156. [CrossRef]
5. Heugues, M. *The Global Emission Game: On the Impact of Strategic Interactions.between Countries on the Existence and the Properties of Nash Equilibria*; Fondazione EEM: Milano, Italy, 2013. Available online: https://www.feem.it/m/publications_pages/NDL2013-108.pdf (accessed on 4 December 2020).
6. Dutta, P.K.; Radner, R. Self-enforcing climate-change treaties. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5174–5179. [CrossRef] [PubMed]
7. Hannam, P.M.; Vasconcelos, V.V.; Levin, S.A.; Pacheco, J.M. Incomplete cooperation and co-benefits: Deepening cli-mate cooperation with a proliferation of small agreements. *Clim. Chang.* **2017**, *144*, 65–79. [CrossRef]
8. Nordhaus, W. Climate Clubs: Overcoming Free-riding in International Climate Policy. *Am. Econ. Rev.* **2015**, *105*, 1339–1370. [CrossRef]
9. Finus, M. Game Theoretic Research on the Design of International Environmental Agreements: Insights, Critical Remarks, and Future Challenges. *Int. Rev. Environ. Resour. Econ.* **2008**, *2*, 29–67. [CrossRef]
10. Carraro, C.; Siniscalco, D. Strategies for the international protection of the environment. *J. Public Econ.* **1993**, *52*, 309–328. [CrossRef]
11. Wood, P.J. Climate Change and Game Theory: A Mathematical Survey, Canberra: The Crawford School of Economics and Government. Available online: http://ageconsearch.tind.io//bitstream/95061/2/Climate%20Change%20and%20Game%20Theory.pdf (accessed on 2 December 2010).
12. North, D.C. *Institutions, Institutional Change and Economic Performance*; Cambridge University Press: Cambridge, UK, 1990.
13. Stiglitz, J.E. *Economics of the Public Sector*; W.W. Norton & Company: New York, NY, USA, 2000.
14. DeCanio, S.J.; Fremstad, A. Game theory and climate diplomacy. *Ecol. Econ.* **2013**, *85*, 177–187. [CrossRef]
15. Barrett, S.; Dannenberg, A. Climate negotiations under scientific uncertainty. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17372–17376. [CrossRef]
16. Barrett, S. Climate treaties and approaching catastrophes. *J. Environ. Econ. Manag.* **2013**, *66*, 235–250. [CrossRef]
17. Mielke, J.; Steudle, G.A. Green Investment and Coordination Failure: An Investors' Perspective. *Ecol. Econ.* **2018**, *150*, 88–95. [CrossRef]
18. Lloyd, W.F. *Two Lectures on the Checks to Population*; Oxford University Press: Oxford, UK, 1833.
19. Hardin, G. The Tragedy of the Commons. *Science* **1968**, *162*, 1243–1248. [CrossRef]
20. Dawes, R. Formal Models of Dilemmas in social decision-making. In *Human Judgment and Decision Processes*; Kaplan, M.F., Schwartz, S., Eds.; Academic: New York, NY, USA, 1975; pp. 88–107.
21. Olson, M. *The Logic of Collective Action. Public Goods and the Theory of Groups*; Harvard University Press: Cambridge, MA, USA; London, UK, 1965.
22. Rapoport, A. Experiments with N-Person Social Traps II: Tragedy of the Commons. *J. Confl. Resolut.* **1988**, *32*, 473–488. [CrossRef]
23. Romagny, B.; Lobry, C.; Canalis-Durand, M. *Tragedy of the Commons and Prisoner's Dilemma*; Document de Travail du Groupement de Recherche en économie Quantitative d'Aix-Marseille (GREQAM), n° 97A20; GREQAM: Marseille, France, 1997; 16p.
24. Diekert, F.K. The Tragedy of the Commons from a Game-Theoretic Perspective. *Sustainability* **2012**, *4*, 1776–1786. [CrossRef]
25. Ragusa, A.; Svelato, A.; Santacroce, C.; Catalano, P.; Notarstefano, V.; Carnevali, O.; Papa, F.; Rongioletti, M.C.A.; Baiocco, F.; Draghi, S.; et al. Plasticenta: First evidence of microplastics in human placenta. *Environ. Intern.* **2021**, *146*, 106274. [CrossRef]
26. Landrigan, P.J.; Stegeman, J.J.; Fleming, L.E.; Allemand, D.; Anderson, D.M.; Backer, L.C.; Brucker-Davis, F.; Chevalier, N.; Corra, L.; Czerucka, D.; et al. Human Health and Ocean Pollution. *Ann. Glob. Health* **2020**, *86*, 151. [CrossRef]
27. Dell, M.; Jones, B.; Olken, B. Climate Change and Economic Growth: Evidence from the Last Half Century. *Am. Econ. J. Macroecon.* **2008**, *4*. [CrossRef]
28. Walker, B.; Barrett, S.; Polasky, S.; Galaz, V.; Folke, C.; Engström, G.; Ackerman, F.; Arrow, K.; Carpenter, S.; Chopra, K.; et al. Looming Global-Scale Failures and Missing Institutions. *Science* **2009**, *325*, 1345–1346. [CrossRef] [PubMed]
29. Bang, G.; Hovi, J.; Sprinz, D.F. US presidents and the failure to ratify multilateral environmental agreements. *Clim. Policy* **2012**, *12*, 755–763. [CrossRef]

30.　Clémençon, R. The Two Sides of the Paris Climate Agreement: Dismal Failure or Historic Breakthrough? *J. Environ. Dev.* **2016**, *25*, 3–24. [CrossRef]

31.　Bauch, C.T.; Earn, D. Vaccination and the theory of games. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13391–13394. [CrossRef] [PubMed]