CHAPTER 14[*]

TRANSFORMING CLIMATE DILEMMAS FROM TRAGEDY TO COOPERATION

BRYAN BRUNS[†]

Abstract: Climate change has often been analyzed as a tragedy of the commons, a social dilemma where cooperation could make everyone better off but incentives induce individuals, businesses, and nations to keep on emitting greenhouse gasses. However, the simple game theory model of tragedy of the commons, Prisoner's Dilemma, is just one of many possible models for climate conflict and cooperation. The topology of payoff swaps in 2x2 games shows relationships between games, including their potential transformations. Changes in the ranking of outcomes can transform Prisoner's Dilemma into a Stag Hunt with the potential for win-win cooperation or Chicken with a shared fear of catastrophe, and then create convergent incentives that yield the best for both in Concord. Models of climate negotiations about whether to abate or pollute can be compactly displayed in a table based on how payoffs from symmetric games combine to form asymmetric games. Maps for transforming climate dilemmas reveal symmetric and asymmetric pathways to climate cooperation through fear of catastrophe, or assuring cooperation that is best for both, or adjusting incentives even if a polluter always want to avoid abating while the other pollutes. Maps for transforming climate games show the diversity of climate dilemmas and potential pathways to cooperation.

Key words: cooperation, climate change, earth governance, global public goods, mechanism design, social dilemmas, topology of 2x2 games

---

[*] This is a submitted manuscript under review for publication in *The Cambridge Handbook of Commons Research Innovations* edited by Sheila Foster and Chrystie Swiney (Cambridge University Press).

[†] Sociologist, Independent Researcher and Consultant. bryanbruns@bryanbruns.com. https://bryanbruns.com

Controlling emissions of greenhouse gasses (GHGs) has often been analyzed as a tragedy of the commons, a social dilemma where cooperation would be better for everyone but incentives lead individuals, businesses, and nations to keep on polluting.[1] This chapter shows how a simple game theory model of tragedy of the commons, Prisoner's Dilemma, is only one of many possible models for climate conflict and cooperation. Changing the ranking of outcomes transforms one game model into another.[2] Mapping these transformations displays pathways for turning tragedy into cooperation.

This chapter first describes how three famous games, Prisoner's Dilemma, Chicken, and Stag Hunt, can transform into each other or into a stable win-win game of Concord. The second section looks at relationships between some previous game theory models for international climate negotiations.[3] The third section presents a more complete map of diversity and potential transformations in climate situations. The fourth section analyzes how the incentive structures in climate situations can change as resource problems worsen and explores pathways for escaping from climate dilemmas by avoiding catastrophe, improving gains from cooperation, or asymmetrically restructuring incentives. The fifth section discusses limitations and extensions of two-person two-choice (2x2) game models for understanding cooperation in coping with climate

---

[1] For example, see Marvin S. Soroos, "Global Change, Environmental Security, and the Prisoner's Dilemma," *Journal of Peace Research* 31, no. 3 (1994): 317–332; Stephen M. Gardiner, "A Perfect Moral Storm: Climate Change, Intergenerational Ethics and the Problem of Moral Corruption," *Environmental Values* 15, no. 3 (2006): 397–413; Gardiner, Stephen M. 2011. *A Perfect Moral Storm: The Ethical Tragedy of Climate Change*. Oxford University Press.; Rony Smead and Ronald Sandler, "Game Theory and the Ethics of Global Climate Change," *Philosophy and Public Issues* 3, no. 1 (2013): 13–23; William Nordhaus, "Climate Clubs: Overcoming Free-Riding in International Climate Policy," *American Economic Review* 105, no. 4 (2015): 1339–70.

[2] David Robinson and David Goforth, *The Topology of the 2x2 Games: A New Periodic Table* (London: Routledge, 2005); Bryan Bruns, "Names for Games: Locating 2 × 2 Games," *Games* 6, no. 4 (October 22, 2015): 495–520.

[3] DeCanio, Stephen J., and Anders Fremstad. "Game Theory and Climate Diplomacy." *Ecological Economics* 85 (2013): 186; Madani, Kaveh. "Modeling International Climate Change Negotiations More Responsibly: Can Highly Simplified Game Theory Models Provide Reliable Policy Insights?" *Ecological Economics* 90 (2013): 68–76. This chapter expands on comments sent to DeCanio and Fremstad about transformations between games, which they acknowledged and discussed in their 2013 journal article.

problems. Overall, this analysis shows that tragic temptation to defect from cooperation is neither inevitable nor the only kind of challenge posed by GHG emissions. A map of how payoff swaps link games displays the diversity of situations, their potential transformations, and multiple pathways to change climate dilemmas from tragedy into cooperation.

FROM SOCIAL DILEMMAS TO CONCORD

The essence of a tragedy of the commons is that pursuing individual interests could lead to collective ruin, even though cooperation would be better for everyone .[4] In Hardin's fable, a farmer who adds another animal to a shared pasture gets all the benefits, while only suffering a fraction of any losses from overgrazing, but if everyone acts this way the pasture itself may be destroyed. The same logic applies to pollution, if each is free to dump their wastes into the environment while others suffer most of the costs. Similarly, if everyone shares in the benefits from reducing GHGs, then potential contributors may be tempted to free ride,[5] enjoying benefits while continuing to pollute.

The game theory story of the Prisoner's Dilemma illustrates a simple model of such conflict between individual and collective interests. Two prisoners are each offered a deal, while held separately and unable to communicate.[6] If one confesses and the other stays silent, then the one who confesses goes free while the other serves a long sentence. If both confess, they receive a moderate sentence. If both keep silent, they still receive a short sentence. Mutual silence would be better than both confessing. However, examining the logic of choices shows that if the other

[4] Garrett Hardin, "The Tragedy of the Commons," Science 162 (1968): 1243–48; Elinor Ostrom, "Tragedy of the Commons," The New Palgrave Dictionary of Economics 2 (2008).
[5] Mancur Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, MA: Harvard University Press, 1971).
[6] A. W. Tucker, "The Mathematics of Tucker: A Sampler. A Two-Person Dilemma: The Prisoner's Dilemma," *The Two-Year College Mathematics Journal* 14, no. 3 (1983): 228–32. Originally presented in a lecture at Stanford University in 1950.

does not confess, it is best to confess. Furthermore, if the other confesses, then it is better to confess. In each case, confessing does better, posing a dilemma for cooperation (at least under the assumed conditions).

The payoff matrix for Prisoner's Dilemma at the top of Figure 1 shows the outcomes from choices to cooperate (C) or defect (D) with preferences for the four outcomes ranked from worst (1) to best (4). In game theory terms, defection is a dominant strategy, better regardless of what the other does. Defection by both leads to a situation from which neither can improve their payoff unilaterally, a Nash equilibrium (shaded). The distinguishing problem in Prisoner's Dilemma is that defection from cooperation (with 3,3 payoffs) leads to an equilibrium with payoffs that are worse for both (2,2) (a Pareto-inferior equilibrium).

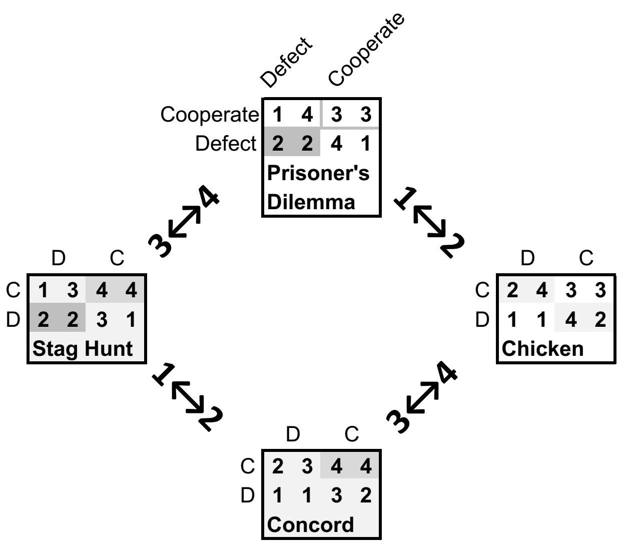

Figure 1: Changes in the ranking of outcomes (payoff swaps) turn Prisoner's Dilemma into Stag Hunt or Chicken, and then into Concord.

However, if both actors defecting from cooperation would lead to the worst outcome (1,1) then the payoff structure is not Prisoner's Dilemma.[7] The game of Chicken takes its name from the story of two cars driving towards each other. If one swerves the other wins, but crashing is worst for both. Both want to avoid disaster. Both swerving is a cooperative second-best outcome. However, rather than cooperating, each would be tempted to defect, seeking to win while the other gets second-worst. If catastrophic climate change could be avoided by only one country reducing (abating) emissions, then each would prefer to get the benefits of abatement by the other without paying the costs. In biology this incentive structure is sometimes called Hawk-Dove. One aggressor could win, but two hawks playing against each other do worst. In politics and climate negotiations, Chicken can model brinksmanship, where two nations try to intimidate or bluff the other into conceding.

Alternatively, cooperation to control GHGs might be best for everyone as long as enough others cooperate, as in a Stag Hunt (assurance) game.[8] Rousseau described a hunter choosing whether to catch a rabbit for sure or getting a stag but only if others cooperate and otherwise getting nothing.[9] A transition from fossil fuels to renewable energy might be similar, where a critical mass of innovation and investment produces the best outcome, as long as enough others

[7] In "Game Theory and Climate Diplomacy" DeCanio and Fremstad repeatedly use the term "catastrophe" for situations in which both polluting is only the second-worst outcome, with abating while the other pollutes as the worst outcome.

[8] This, arguably, is the case for controlling gasses that damage the ozone layer, as DeCanio and Fremstad suggest in "Game Theory and Climate Diplomacy."

[9] Jean-Jacques Rousseau, *A Discourse upon the Origin and the Foundation of the Inequality among Mankind*, 2004 https://www.gutenberg.org/ebooks/11136; Brian Skyrms, *The Stag Hunt and the Evolution of Social Structure* (Cambridge University Press, 2004); A. K Sen, "Isolation, Assurance and the Social Rate of Discount," *The Quarterly Journal of Economics* 81, no. 1 (1967): 112–124; Carlisle Ford Runge, "Common Property Externalities: Isolation, Assurance, and Resource Depletion in a Traditional Grazing Context," *American Journal of Agricultural Economics* 63, no. 4 (1981): 595–606; Cole, Daniel H., and Peter Z. Grossman. "Institutions Matter! Why the Herder Problem Is Not a Prisoner's Dilemma." *Theory and Decision* 69, no. 2 (2010): 219–231.

also cooperate. The crucial distinction between Prisoner's Dilemma and Stag Hunt is: assuming everyone else cooperates, is there still a temptation to defect?[10]

In the game of Concord,[11] incentives align so cooperation is better regardless of what the other does. If both players cooperate, both get their best payoff (4,4). Recommendations for carbon fees and abolishing subsidies that distort prices for fossil fuels aim to have the "invisible hand" of market prices lead individual decisions to an outcome that is best for all.[12] Some research suggests that investing in efficiency and renewable energy is worthwhile even at current (distorted) prices,[13] in which case reducing GHG emissions might already be the best strategy without depending on how others act. Redesigning carbon fee policies to better attract voters could further improve incentives to reduce emissions.[14]

Changes can swap the ranking of outcomes and transform one game into another, as Figure 1 shows. Swapping the two lowest-ranked outcomes (1↔2) turns Prisoner's Dilemma into Chicken. Swapping the two highest-ranked outcomes (3↔4) turns Prisoner's Dilemma into Stag Hunt. Swapping the top-ranked outcomes in Chicken to make cooperation more rewarding turns Chicken into Concord; Swapping the lowest-ranked outcomes in Stag Hunt so that cooperation also minimizes risks turns Stag Hunt into Concord.

For understanding climate problems and solutions, crucial questions concern the current incentive structure and how that could be changed. Are we in a situation where both failing to cooperate leads to the worst outcome as in Chicken, or trapped in a potential tragedy like

[10] Amadae, S. M. *Prisoners of Reason: Game Theory and Neoliberal Political Economy*. Cambridge University Press, 2016.
[11] Names are from Bruns, "Names for Games." Robinson and Goforth used the name "No Conflict."
[12] Nordhaus "Climate Clubs."
[13] Lovins, Amory. *Reinventing Fire: Bold Business Solutions for the New Energy Era*. Chelsea Green Publishing, 2013.
[14] Klenert, David, Linus Mattauch, Emmanuel Combet, Ottmar Edenhofer, Cameron Hepburn, Ryan Rafaty, and Nicholas Stern. "Making Carbon Pricing Work for Citizens." *Nature Climate Change* 8, no. 8 (2018): 669–677.

Prisoner's Dilemma? Could we control emissions in a way that would be best for everyone? The climate literature has sometimes recognized the potential for transformation between games, but usually only for change between a few neighboring games.[15]

CLIMATE NEGOTIATIONS: ABATE OR POLLUTE?

DeCanio and Fremdstad explored 2x2 game models for climate negotiations.[16] They restricted their analysis to games where both abating was better than both polluting and where polluting rather than abating always made the other worse off. They examined all possible 2x2 games with a strict rank ordering of outcomes. Out of 144 order games, DeCanio and Fremstad found twenty-five relevant to climate negotiations.

Five of the climate games were symmetric, where the two actors each face the same pattern of payoffs, including Prisoner's Dilemma, Chicken, Stag Hunt, and Concord.[17] Swapping the ranking of the middle outcomes (2↔3) turns Concord into Harmony, where dominant strategies also lead to win-win.

Although DeCanio and Fremstad used Robinson and Goforth's "New Periodic Table" to identify 2x2 climate games,[18] they presented their results as a series of separate tables for different groups of games. Figure 2 uses the structure of payoff swaps in the "New Periodic

[15] See, for example, Jason K. Levy, Keith W. Hipel, and N. Howard, "Advances in Drama Theory for Managing Global Hazards and Disasters. Part II: Coping with Global Climate Change and Environmental Catastrophe," *Group Decision and Negotiation* 18, no. 4 (2009): 317–334; Karen Pittel and Dirk TG Rübbelke, "Transitions in the Negotiations on Climate Change: From Prisoner's Dilemma to Chicken and beyond," *International Environmental Agreements: Politics, Law and Economics* 12, no. 1 (2012): 23–39. Barrett, Scott. "Collective Action to Avoid Catastrophe: When Countries Succeed, When They Fail, and Why." *Global Policy* 7 (2016): 45–55.
[16] "Game Theory and Climate Diplomacy."
[17] Under some conditions, a competitive person might "defect" in Concord to try to get the outcome where they do better (3,2), for example if they could choose first so the other has only a choice between worst and second-worst. Due to this susceptibility to a competitive concern to maximize the difference in payoffs, some social psychologists call this payoff structure Max-Diff. See Kelley, Harold H., and John W. Thibaut. *Interpersonal Relations: A Theory of Interdependence*. John Wiley & Sons Inc, 1978. If nations pursue relative advantage, not absolute gains, this could pose an issue in Concord-type situations, especially if one actor could take initiative and commit themselves.
[18] Robinson and Goforth, "Topology of 2x2 Games."

Table" to display the five symmetric games along a diagonal axis. Their payoffs combine to form the twenty asymmetric games, compactly displaying the relationships among climate games in a single table.[19]

The twenty asymmetric games are composed of ten pairs on either side of the diagonal that differ by exchanging positions for the Row and Column players (reflections). In the lower left of Figure 2, Harmony, Concord, and three pairs of asymmetric games have dominant strategies for one or both that lead to a single "no conflict" equilibrium where both abate. In the center, Prisoner's Dilemma and a pair of asymmetric neighbors have an inefficient "tragic" equilibrium where both pollute. In five pairs of "unhappy" asymmetric games, one pollutes while the other abates and does worse. Another pair of games have no equilibrium (in pure strategies) and a cyclic incentive structure where one or the other always wants to change their move. However, the 4,3 outcome could be a prominent focal point[20] for agreeing to abate.

[19] The symmetric games also act as coordinates to name and locate the asymmetric games, Bruns, "Names for Games." Figures 2 and 3 also show layer, row, and column numbers that compose Robinson and Goforth's game index numbers, as used by DeCanio and Fremstad and by Madani.
[20] Schelling, T. C. *The Strategy of Conflict*. Cambridge MA: Harvard University Press, 1960.

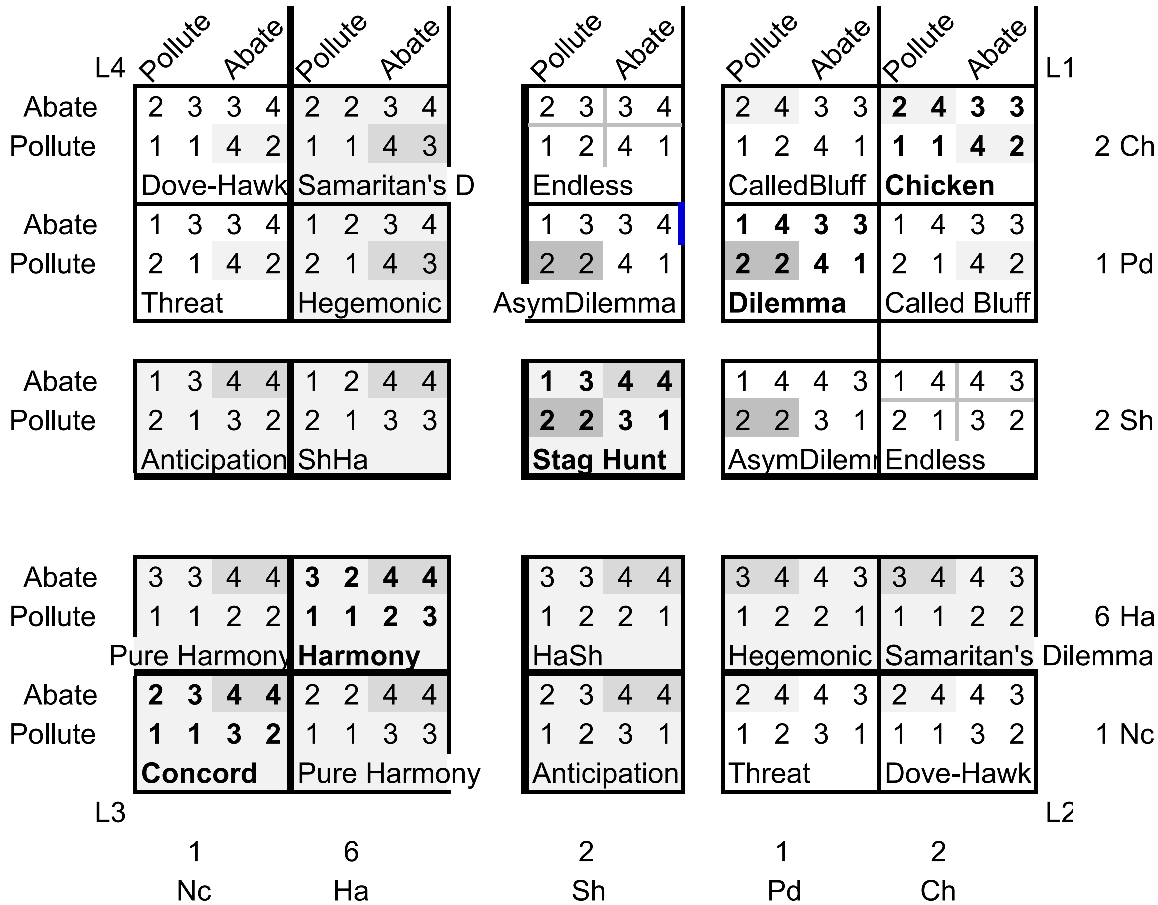

Figure 2. Climate negotiation games model choices to abate or pollute. Symmetric games form a diagonal from lower left to upper right. Their payoffs combine to form asymmetric games.

A Map for Changing Games

The set of five symmetric games on the diagonal in Figure 2 extends to include five more games, coordination, second-best, and harmonious climate situations that DeCanio and Fremstad's restrictions excluded. Payoffs from the twelve symmetric games combine to form asymmetric games, making a more complete map of transformations between situations. This section describes the additional symmetric games, and the "Periodic Table of 2x2 Games" in more detail.

In Chicken, one or the other does much better at the rival equilibria, best versus second-worst (4,2). The rival equilibria in the neighboring games of Hero and Leader (Battle) are less unequal, best versus second-best (4,3).[21] They could model situations where it would be more efficient for only one to abate. Hero could model where both abating would be too much, as might occur with geoengineering. Other choices could also require coordination but be mutually exclusive and offer advantages to one nation or the other (or group of countries, such as fossil fuel exporters versus importers). One climate debate concerns directly switching to renewable energy versus using natural gas during a transition. Everyone wants to avoid the worst outcome of climate catastrophe and the second-worst outcome of uncoordinated and inefficient investment. However, countries may differ over which path to take.

Stag Hunt has one win-win equilibrium and an inferior equilibrium where both get second worst (2,2). By contrast, Assurance and (Strict) Coordination have an inferior equilibrium where both get second-best (3,3) . These two games could model a situation where only concerted strategies can succeed and such an effort by only one would make the other worse off (unlike DeCanio and Fremstad's assumptions). For two big developing countries the only feasible strategy might be to first rely on fossil fuels, temporarily worsening global warming, to reach a level of economic development where they both could switch to renewable energy and carbon capture. Only a joint effort might be sufficient to gain synergies or cross a threshold for success. If such a strategy by one big developing country alone were at least able to avoid a catastrophic outcome for that country then the situation would be like Strict Coordination. If such a solo strategy led to that country getting their worst outcome, then the situation would be like Assurance. These six games all pose coordination problems of selecting between alternative

[21] Gardiner's discussion of "battle of the sexes" mainly concerns a dynamic coordination problem of assembling an effective coalition, not rivalry over which of two mutually desirable equilibria to select.

equilibria, where failure to successfully coordinate would be worse for both (a Pareto-inferior outcome).

In the games of Deadlock and Compromise, dominant strategies lead to second-best, differing only by the location of the two lowest-ranked payoffs. As long as impacts are not severe, some level of GHG emission may be tolerable, even if not ideal. This could be the case if abatement was not economically worthwhile or if each nation would most prefer to pollute while the other abates. However, if mutual abatement would be better than both continuing to pollute, that would swap the middle-ranked outcomes and convert the payoff structure from Deadlock into Prisoner's Dilemma.

Middle swaps turn Harmony into Peace, where dominant strategies also converge on win-win. This game could model a situation where one side's choice to cooperate makes things worse for the other unless they also cooperate. For example, a climate treaty between nations could impose trade restrictions that penalize those continuing to pollute. Contrary to DeCanio and Fremstad's assumptions (restrictions), there are relevant climate situations where abatement by one might be better than by both or where abatement by one could make the other worse off unless they also abate.

Figure 3 displays Robinson and Goforth's "Periodic Table of 2x2 Games" where payoff swaps link neighboring games.[22] The twelve strict symmetric ordinal games discussed above form a diagonal axis. Their payoffs combine to make asymmetric games, elegantly revealing relationships between 2x2 games.

In the detailed structure of the topology, low swaps (1↔2) form "tiles" of four games. Middle swaps (2↔3) link tiles within four "layers." In the display, each layer wraps top-to-

[22] Robinson and Goforth, *Topology of 2x2 Games*. Bruns, "Names for Games."

bottom and side-to-side like a doughnut (torus). In each layer, the top two payoffs have the same alignment: in a win-win cell as in Harmony; diagonally opposed as in Prisoner's Dilemma and Chicken; or in the same row or column. High swaps (3↔4) connect games across layers and link the entire table top-to-bottom and side-to-side. Thus at the center of the table, high swaps change Prisoner's Dilemma into an Asymmetric Dilemma and then Stag Hunt. High swaps for one or the other also transform the coordination problems in Strict Coordination and Assurance into rivalries on the Layer 1 tile with Hero and Leader. High swaps for both transform Deadlock into Peace and Compromise into Harmony.[23]

[23] To visualize more high swaps, horizontally switch the Harmony row of six games with the equivalent row on Layer 2; high swaps for row payoffs then transform into the game in the row above. Harmony turns into *PcDl* and Samaritan's Dilemma, *HaCh*, turns into Pure Samaritan, *PcSh*. High swaps for row payoffs similarly link the rows with Deadlock and Compromise. Similar column swaps slide vertically. Rows (and columns) for coordination and cyclic tiles cross diagonally between layers.

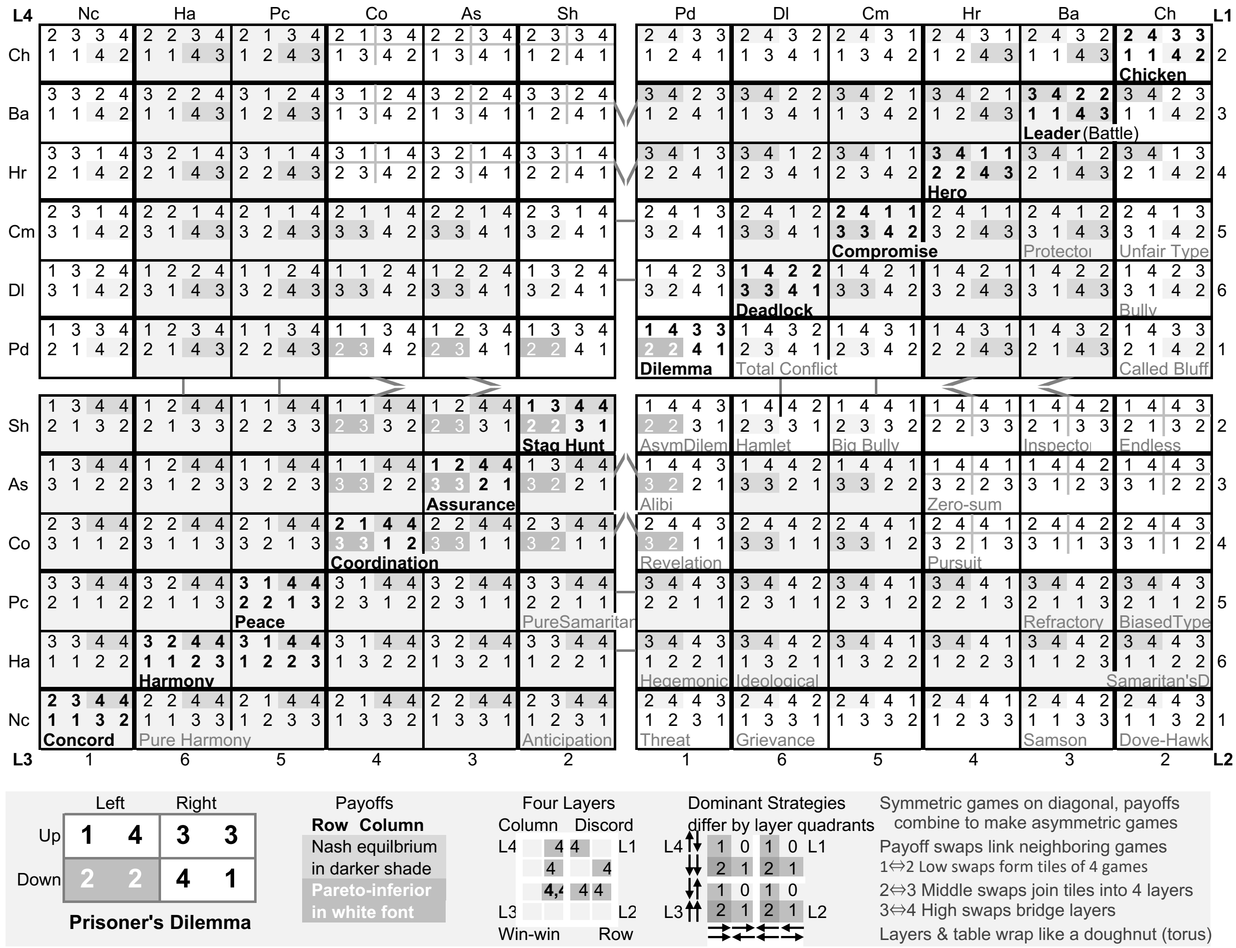

Figure 3. The topology of 2x2 games shows potential transformations. Payoff swaps link neighboring games. Symmetric games lie on a diagonal axis; their payoffs combine to form asymmetric games. High swap transformations link layers.

PATHWAYS FOR TRANSFORMATION

As greenhouse gasses accumulate, impacts worsen, and gains from abatement increase, the situation could move through a series of games.[24] This can be visualized in transformations

[24] Kaveh Madani, "Game Theory and Water Resources," *Journal of Hydrology* 381 (2010): 225–238; Madani, "Modeling Climate Change;" Bora Ristić and Kaveh Madani. "A Game Theory Warning to Blind Drivers Playing Chicken with Public Goods." *Water Resources Research* 55, no. 3 (2019): 2000–2013. Madani and Ristić use

between games in Figures 2 and 3, looking first at symmetric transformations along the diagonal axis. Initially, each nation might prefer to pollute while the other abates but would settle for second-best where both pollute, in which case they would start in Deadlock. As climate impacts become more severe, abatement may be worthwhile if both cooperate, becoming a Prisoner's Dilemma, where each would be tempted to defect, trying to gain the benefits of the other's efforts without paying the costs. As conditions worsen the situation could turn into Chicken, where both polluting is worst for both.[25] Cooperation would be second-best, but each would still be tempted to defect from cooperation so they could keep polluting while the other abates enough to avoid catastrophe. As conditions further deteriorate, each might have sufficient incentive to act alone to mitigate disaster regardless of what the other does. This would be like Concord, where dominant strategies converge on abatement. However, by then it might be long past climate tipping points that lead to catastrophic changes.[26]

Outcomes of climate situations, and their rankings, could change for many reasons. Improvements in technology, institutions, and financing, such as for renewable energy, could make cooperation more efficient and profitable. Governments, businesses, and other organizations could develop shared strategies, support norms favoring a particular outcome, or establish rules that regulate behavior and incentives.[27] Expectations could shift, for example,

---

formulas and cost curves to model changes between payoff structures, which this chapter visualizes for a broader range of transformations between games.

[25] If costs are asymmetric, then one will have a Chicken payoff sooner, making a game of Called Bluff. One keeps polluting, while the other starts abating, incurring costs and doing worse. In Figure 8 in Ristić and Madani "Game Theory Warning" this is the red zone on the asymmetric path from Prisoner's Dilemma to Chicken.

[26] Barrett, Scott, and Astrid Dannenberg. "Sensitivity of Collective Action to Uncertainty about Climate Tipping Points." *Nature Climate Change* 4, no. 1 (2014): 36–39. Barfuss, Wolfram, Jonathan F. Donges, Vítor V. Vasconcelos, Jürgen Kurths, and Simon A. Levin. "Caring for the Future Can Turn Tragedy into Comedy for Long-Term Collective Action under Risk of Collapse." *Proceedings of the National Academy of Sciences*, 2020

[27] Ostrom, Elinor. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press, 2005.

based on new information or greater fear of risks.[28] The ranking of outcomes could also change due to sympathy, more concern for what happens to others or a change in moral judgements. If the payoff pattern changes only for one player, the situation becomes asymmetric, a possibility that can help think about more options for escaping climate dilemmas.

Depending on the starting point, multiple pathways may reach climate cooperation. The solution for Prisoner's Dilemma is often to turn it into a Stag Hunt, where cooperation is best for both.[29] A governance solution, unsuccessful so far, would be an enforceable international agreement imposing penalties or rewards that overcome temptation to defect. Concern for others, including all humans alive now, future generations, and nature, should transform the evaluation of outcomes so that cooperation is the only morally acceptable choice.[30] Or, as mentioned above, technological changes and coordinated investments might suffice to make abatement best for all.

If both not cooperating would be the worst outcome, that poses a problem like Chicken. If the benefits of cooperation can be increased and temptation to defect reduced symmetrically, that turns Chicken into Concord. Abatement by one nation or other actor, asymmetrically, might initially lead them to do worse than the other, as in the game of Dove versus Hawk. However, this inequality might be accepted as a necessary "strategic loss."[31] This could facilitate learning, negotiation, and further transition.[32]

[28] Changing the perceived ranking of outcomes creates a different "effective matrix." Harold H. Kelley and John W. Thibaut, *Interpersonal Relations: A Theory of Interdependence* (John Wiley & Sons Inc, 1978); Kollock, Peter. "Social Dilemmas: The Anatomy of Cooperation." *Annual Review of Sociology* 24, no. 1 (1998): 183–214; Kollock, Peter. "Transforming Social Dilemmas: Group Identity and Cooperation." *Modeling Rational and Moral Agents*, 1998, 186–210.

[29] Skyrms, *The Stag Hunt*.

[30] Gardiner, *Perfect Moral Storm.* The chapter by Hudson in this volume examines challenges that time poses for commons governance and intergenerational equity.

[31] Ristić and Madani, "A Game Theory Warning."

[32] Dubash, Navroz K. "Revisiting Climate Ambition: The Case for Prioritizing Current Action over Future Intent." *Wiley Interdisciplinary Reviews: Climate Change* 11, no. 1 (2020): e622.

Even if one nation (or set of nations) never believes both polluting would be the worst outcome, there might still be ways to reach a situation where incentives lead both sides to abate. If one nation thinks abatement is worth doing even without cooperation, that changes their payoff pattern to Chicken, changing the game from Prisoner's Dilemma to Called Bluff, where they abate and get second-worst. If that nation comes to believe that mutual abatement would be the best outcome (rather than preferring to free ride if the other abates), then their payoff becomes that in Concord and the game would become Threat.[33] They might decide abatement is a moral imperative, or raise their expectations about the mutual benefits of a sustainable economy based on renewable energy, or see abatement it as a necessary transitional step that will be followed by others. The other nation might continue to pollute and get their "best" outcome. However, Threat is susceptible to transformation. The abater could threaten to resume polluting, a veto that would block the other nation from getting best or second-best. If the other changes to rank abatement as better than polluting, for whatever reasons, including information, persuasion, moral reconsideration, reputational risks, side payments, linkage, sanctions, or other pressures, that would reach a stable win-win game of Anticipation. This pathway illustrates how assuming symmetry or conditional cooperation only when others cooperate could blind analysts and decisionmakers to less symmetrical, less ideal, and less equitable but potentially feasible pathways to solutions.

Figure 5 shows the games formed by combining payoffs from Chicken, Prisoner's Dilemma, Stag Hunt, and Concord. This makes a simpler microcosm (subspace) of possible games showing the three pathways for 1) turning Prisoner's Dilemma into Stag Hunt at the center of the diagram and then into Concord, 2) converting Chicken into Concord (perhaps via

[33] Guyer, Melvin J., and Anatol Rapoport. "Threat in a Two-Person Game." *Journal of Experimental Social Psychology* 6, no. 1 (1970): 11–25.

Dove-Hawk), and 3) transforming asymmetrically from Prisoner's Dilemma through Called Bluff to Threat and win-win cooperation in Anticipation.

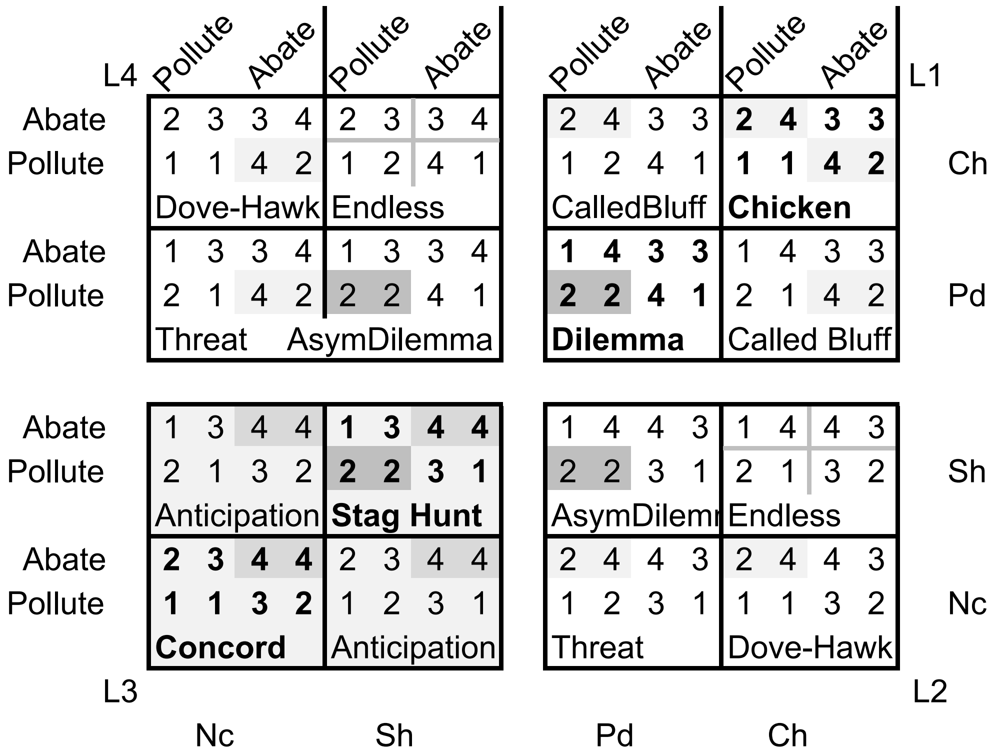

Figure 4. In a microcosm for transforming from pollution to abatement, changes that swap the ranking of outcomes turn tragic conflict in Prisoner's Dilemma into win-win cooperation.

LIMITATIONS AND EXTENSIONS OF SIMPLE 2X2 MODELS

Game theory models offer insights into climate problems, but have many limitations,[34] especially if models make assumptions typical of early mathematical game theory, such as no communication, single-shot interaction, simultaneous decisions, irreversible moves, complete

[34] Madani, "Modeling International Climate Change."

information, payoffs that already incorporate other-regarding preferences, risk-neutrality, no external enforcement, no side payments, and no links to other situations. Conversely, research has shown how different conditions can contribute to cooperation, even in Prisoner's Dilemma and multi-person dilemmas.[35] Cooperation may be more successful with repeated interaction, communication, and incremental commitment. Most people are inclined to cooperate and willing to cooperate as long as others do so. Most care about what happens to others and are willing to punish those who do not cooperate. People are not prisoners, and often can "change the game" by making rules and other institutions that enable them to cooperate, including sustainably governing commons.[36]

Promising opportunities for further analyzing transformations between climate dilemmas include systematically applying multiple models beyond Prisoner's Dilemma and exploring asymmetric pathways where unequal and unsatisfactory outcomes could encourage efforts to find more equitable and mutually acceptable results. A broader menu of models may also help to understand the diversity of interconnected action situations between and within nations.[37]

---

[35] Some useful sources in a vast literature include Rapoport, Anatol, Melvin J. Guyer, and David G. Gordon. The 2 x 2 Game. Univ of Michigan Press, 1976; Axelrod, Robert. The Evolution of Cooperation. New York: Basic Books, 1984; Kollock, Peter. "Social Dilemmas: The Anatomy of Cooperation." Annual Review of Sociology 24, no. 1 (1998): 183–214; Nowak, Martin A. "Five Rules for the Evolution of Cooperation." Science 314, no. 5805 (2006): 1560–1563; Taylor, Michael. Rationality and the Ideology of Disconnection. Cambridge University Press, 2006.

[36] Ostrom, Elinor. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press, 1990. Ostrom, Elinor, Roy Gardner, and James Walker. 1994. *Rules, Games and Common-Pool Resources*. Ann Arbor: University of Michigan Press.

[37] Putnam, Robert D. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42, no. 3 (ed 1988): 427–60. Aklin, Michaël, and Matto Mildenberger. "Prisoners of the Wrong Dilemma: Why Distributive Conflict, Not Collective Action, Characterizes the Politics of Climate Change." *SSRN*, 2018. Ostrom, Elinor. "A Polycentric Approach for Coping with Climate Change." The World Bank, 2009; Daniel H. Cole, "Advantages of a Polycentric Approach to Climate Change Policy," *Nature Climate* Change 5, no. 2 (2015): 114–118; Jordan, Andrew, Dave Huitema, Harro van Asselt, and Johanna Forster. *Governing Climate Change: Polycentricity in Action?* Cambridge University Press, 2018.

CONCLUSIONS: TRANSFORMING CLIMATE DILEMMAS

Changes in the ranking of outcomes show how the tragic temptations of Prisoner's Dilemma can turn into potential coordination in Stag Hunt or shared fear of a worst outcome in Chicken, and then transform into convergent incentives to win-win in Concord. Combining payoffs from symmetric games to form asymmetric games reveals relationships between models of climate negotiations about whether to abate or pollute. The topology of payoff swaps in 2x2 games maps potential transformations and shows symmetric and asymmetric pathways to climate cooperation through fear of catastrophe if both pollute, changes that make cooperation best for both, or restructuring incentives even if a polluter always sees abating while the other pollutes as the worst outcome. A map for changing games displays the diversity of climate dilemmas and potential pathways to cooperation.