

# Prisoner's Dilemma and Climate Change Mitigation

Andy Sheu

University of California, Berkeley

April 25, 2018

## Abstract

People pondering the real world implications of the prisoner's dilemma have often considered climate change mitigation to be one such scenario. The prisoner's dilemma and our perception of climate change are so similar that the oversimplifications and flaws inherent to the prisoner's dilemma also apply to how we think of global warming. Experiments using more accurate variants of the prisoner's dilemma reveal a course of action different from our current strategies. This paper first examines how the flaws in a traditional prisoner's dilemma model of climate change mitigation mirrors flaws in our perception of climate change and ultimately discusses how those flaws result in mistakes in climate negotiation strategies.

## Background

The prisoner's dilemma is a famous game theory scenario wherein two people are arrested. The police have evidence to convict both of a minor crime, but the police suspect that they jointly committed a more serious crime, of which the police have insufficient evidence. The two prisoners are separated into different rooms and are each offered a deal with two options: stay silent (cooperate) or testify against the other (betray). The resulting outcomes (in years of prison time) are listed below in Figure 1, with the numbers colored in accordance to the color of the respective person's choices.

	C	B
C	1 / 1	0 / 3
B	3 / 0	2 / 2

Figure 1. Prisoner's dilemma payouts (arbitrary numerical assignment)

The crucial concept behind the prisoner's dilemma is the fact that each person is better off betraying, resulting in a game state where both players are worse off. If the other person

cooperates, switching from cooperating to betraying reduces your sentence from 1 year to nothing. If the other person betrays, you reduce your sentence from 3 years to 2 years. Notice that, either way, you reduce your own sentence by betraying. Yet in each attempting to better their own positions, both players paradoxically end up worse off than if they had both cooperated.

### **Introduction**

In attempting to simplify the reality of climate change mitigation into a researchable model, many (including Abrams and Newkirk, to an extent) have likened climate change to the prisoner's dilemma. At first glance, this comparison seems to be perfectly analogous. In fact, the prisoner's dilemma is precisely what we fundamentally perceive climate negotiation to be; each country has an enormous economic incentive to defect and free-ride off of others' sacrifices, ultimately driving the system towards a state of equilibrium where all parties are worse off. But therein lies the problem: the prisoner's dilemma is very much dissimilar from reality. Therefore, the flaws in the prisoner's dilemma reflect the flaws in our own perception of climate change mitigation.

Various researchers have attempted to devise variants of the prisoner's dilemma to address those underlying flaws and construct a more accurate model. They have introduced a variety of complexities including iteration (Wood, Ostrom, Irwin), reputation (Wood, Ostrom, Irwin), thresholds (Tavoni, Iris, Dannenberg), inherent inequities (Tavoni), ambiguity (Dannenberg, Ostrom), elected delegates (Tavoni, Iris), interactions (Dannenberg, Wood, Irwin), burden distribution (Wood, Aaron K, Page, Irwin), and communication (Tavoni). Experiments involving these additional intricacies reveal a myriad of previously unapparent actions and

factors that are strongly associated with success in reaching thresholds. Given that these strategies are developed from models that attempt to address flaws in how we understand climate change mitigation and that many of these strategies run contrary to how we have attempted to address climate change, we need to consider modifying our approach to climate change.

In combination, these two points about the relationship between climate change strategy and the prisoner's dilemma form the crux of my argument. Firstly, the flaws in the prisoner's dilemma model mirror the flaws in our perception of climate change mitigation. Moreover, strategies that succeed in prisoner's dilemma models which fix the aforementioned flaws demonstrate a course of action that wildly deviates from our current path. Ultimately, though game theory and the prisoner's dilemma are recklessly simplistic and overly theoretical topics that contrast sharply with the complexities and realism of climate change mitigation, game theory still offers an array of insights into how we ought to reconceptualize climate change and modify strategies for lessening its impacts.

### **Prisoner's Dilemma Flaws**

#### **Model Simplifications**

Although the prisoner's dilemma bears a strong resemblance to climate change mitigation, it suffers from the oversimplification inherent to modeling real world occurrences in generalized theoretical models. For one, it models climate negotiations as a two person interaction. This simplification mirrors how we Americans tend to think about climate agreements mostly with respect to our economic competition with China. In reality, climate negotiations often involve hundreds of nations. The Paris Climate Accord, for instance, had 195

signatories and 176 parties<sup>1</sup> while the Copenhagen Accord was agreed upon by 114 nations<sup>2</sup>. In reducing climate change to an issue between the world's two most developed countries, we risk alienating potential allies and overlooking the very real and disastrous consequences global warming has on many smaller nations. Along those lines, climate change is also very dissymmetrical: it is not an interaction among equals. As Evan Abram asserts, countries all face different consequences and possess differing degrees of efficacy.<sup>3</sup> For example, "the defection of just one...major [party] would likely kill any deal. Yet some smaller, less industrialized nations would add almost nothing to emissions reductions."<sup>4</sup> Glenn Althor's research concurs, noting vast discrepancies between nations in both climate change vulnerability and emission levels.<sup>5</sup> This disparity comes into play both in distributing the burden of climate change and understanding how decisions of a select group of nations overshadows the decisions of the rest of the world. For the most part, the fate of people oceans away lies in our hands, not theirs. In choosing our path and discussing how others should choose theirs, we need to keep this in mind.

Even though America under the Trump administration has chosen its national climate policy, signalling its intent to withdraw from the Paris Climate Agreement and to continue to emit carbon for economic prosperity, our nation can still change its mind and contribute to the fight against global warming. This is symbolic of another flaw: climate change mitigation is a

---

<sup>1</sup> "UN, United Nations, UN Treaties, Treaties." United Nations.

<sup>2</sup> "Information Provided by Parties to the Convention Relating to the Copenhagen Accord." UNFCCC.

<sup>3</sup> Abrams, Evan. "Climate Change Is the Ultimate Prisoners' Dilemma | GRI." Global Risk Insights.

<sup>4</sup> Abrams.

<sup>5</sup> Althor, Glenn, James E. M. Watson, and Richard A. Fuller. "Global Mismatch between Greenhouse Gas Emissions and the Burden of Climate Change." Nature News.

series of repeated interactions, rather than just a one-time decision.<sup>6</sup> Furthermore, each decision has long-lasting effects, both in the severity of global warming and in each country's circumstances. As a result, interactions in the present cascade down and greatly reshape future interactions. One such effect that we often overlook is the impact our policies have on our international reputation. Refusing to act when other nations are in dire peril reflects poorly on us as a nation and could very feasibly contribute to a future scenario in which other nations refuse to help us. Moreover, our inaction could possibly cause other nations, including China, to also withdraw in retaliation. If that were to happen, the global fight against climate change would utterly collapse.

Thus far, political entities have framed climate change as binary choice: take action against climate change or don't. Likewise, the prisoner's dilemma offers each player two distinct choices: cooperate or defect. But this framing simply isn't accurate because climate change mitigation is a question of what, how much, and when.<sup>7</sup> We have a slough of potential actions available to us, each of whom have wildly different costs and benefits. Lumping all possible solutions into one vague category muddies our ability to adequately predict future emissions and necessary reductions. Along the idea of necessary reductions, we have set emission targets to try to reach arbitrary thresholds and have focused on meeting specific numerical objective. Yet, we follow that path even though we aren't entirely sure what those numbers mean. In doing so, we have committed to pursuing arbitrary quantitative goals rather than trying to do as much as we can to limit the effects of climate change.

---

<sup>6</sup> Spring, Judith. "Why Climate Change Is Essentially Not a Prisoner's Dilemma." IMSD International Master in Sustainable Development and Corporate Responsibility RSS.

<sup>7</sup> Wood, Peter John. "Climate Change and Game Theory: A Mathematical Survey." Environmental Economics Research Hub Research Reports, 7.

## Model Instabilities

The other major issue with the prisoner's dilemma model is that it fails to capture than uncertainties and fluctuations inherent to modeling future realities. For one, the payouts are largely unknown. As Abrams puts it, "it is not clear...what the costs of economic adjustment would be under a climate agreement and how these costs would be distributed" and even total cooperation may be insufficient to avert the most significant effects of global warming.<sup>8</sup> In other words, we are unsure about how mitigation will affect the economy and which consequences we can even avoid. Similarly, the payout model changes over time as past actions modify the present, technologies innovate, and the world randomly fluctuates. For example, Mark Sagoff suggests that climate change is such a long-term problem that our choices now will have their largest impacts on unborn generations and within future markets which we cannot even begin to conceive.<sup>9</sup> Put simply, we do not understand the long-term economic implications of our actions. For instance, it is entirely possible that reducing fossil fuel dependency and investing in cleaner energy sources could end up being a fiscally beneficial, rather than harmful, action. Furthermore, as suggested by Judith Spring and Mark Sagoff, inaction or insufficient emission reductions could conceivably introduce a plethora of unknown long-term global issues such as water wars, reduced trade, or refugees.<sup>10 11</sup> Given that these developments have yet to occur and that we have no way of predicting them, we simply aren't able to accurately calculate the aggregate economic costs of investing in climate change mitigation. As a result of the uncertainty revolving around the consequences and costs of our actions, you cannot safely assume betraying to be more

---

<sup>8</sup> Abrams.

<sup>9</sup> Sagoff, Mark. "Climate Change Is Not a Tragedy of the Commons." Learn Liberty.

<sup>10</sup> Spring.

<sup>11</sup> Sagoff.

advantageous than cooperating, just as you cannot safely assume emissions reductions to be an economic sacrifice in the long run.

## **Optimal Strategies**

### Thresholds and Leadership

One experimental variant of the prisoner's dilemma that attempts to address the flaws in the original is a variant in which a large group of people (more than two) are all given an arbitrary amount of starting money and get to keep their remaining money if the group collectively reaches a specified donation threshold over an unknown number of rounds. In practice, scientists used slightly modified versions to derive a control and experimental group, introducing changes such as unknown thresholds or communication. One such experiment had groups electing delegates who receive messages from their group and act on behalf of the group.

<sup>12</sup> Doruk Iris discovered that delegation and messaging decreased contributions to the public good and that “delegates [tended] to focus on the lowest contribution level suggested by non-delegates.”<sup>13</sup> In other words, elected representatives, especially when dependent on reelection, gravitate towards the least ambitious actions. Knowing this, we ought to reform how we represent our nation in climate diplomacy to more accurately reflect the aggregation of our collective beliefs, rather than concentrating on the most inertial suggestions. Although there are no readily apparent solutions, we can try to reduce its effects by publically emphasizing climate change policies when electing officials.

---

<sup>12</sup> Iris, Doruk, Jungmin Lee, and Alessandro Tavoni. "Delegation and Public Pressure in a Threshold Public Goods Game: Theory and Experimental Evidence." Grantham Research Institute on Climate Change and the Environment, 17.

<sup>13</sup> Iris, 24.



In addition to our elected officials, modified prisoner's dilemma experiments show us that the rich are also important leaders in combating climate change. In an experiment where players started with different amounts of starting currency, Alessandro Tavoni asserts that "leadership appears to be an important engine of collective action, as successful groups tended to eliminate inequality over the course of the game [and that] rich players signalling willingness to redistribute their funds early on in the game proved to be a precursor for success."<sup>14</sup> Thus, the rich are a crucial factor in determining our ability to successfully address climate change. Understanding that assistance from the wealthy is imperative in the fight against global warming, we need to understand and communicate why the rich, especially, should contribute. Firstly, emissions reductions and investment in green technology are beneficial economically in the long term. Fossil fuels will become increasingly expensive in the next few decades, whether through the pressures of climate change or their radically dwindling supply. Countries and companies need to reduce their dependence on fossil fuels in order to avoid economic collapse when fossil fuels are inevitably used up. China, for instance, has recognized this fact and has invested in cleaner energy to prepare for the eventuality.

Additionally, climate change hurts international trade. In the modern globalist economy, the collapse of one nation has serious ramifications on other nations. It simply isn't accurate to say that climate change isn't our problem. Even if we aren't directly harmed by its effects, we will suffer indirectly through things like increased costs, wars, fleeing refugees, and isolationism. Moreover, climate change inaction will likely end up being more expensive than climate change prevention as changing times necessitate increased taxes and reduced business as we need to

---

<sup>14</sup> Tavoni, Alessandro. "COP21: Can Game Theory Predict a Global Deal on Climate Change?" Grantham Research Institute on Climate Change and the Environment.

devote an increasing amount of money towards disaster relief. Furthermore, reduced biodiversity will prevent certain medicines from ever being developed and rising temperatures and sea levels will result in reduced resource access and increased costs globally. In the end, avoidance of unexpected costs could very well exceed the costs of carbon emission reductions.

Ultimately, the biggest reason that the rich should cooperate in climate change mitigation, according to Edward Page, is because “this is the most efficient way of achieving the desired outcome of greater equality”<sup>15</sup> under which the “benefits and burdens [of climate change] should be distributed so that inequality is minimised.”<sup>16</sup> Not only does action among the wealthy have the greatest potential emissions reductions given their overwhelming power in corporate structures, but the choices of these individuals are the ones that are most likely to inspire change in everyday people, domestic legislation, and international climate policy. Combined with the fact that they are the demographic that sacrifices the least, relatively speaking, the rich are the group most capable of sparking the change necessary to avert our generation’s defining global crisis.

### International Climate Diplomacy

In addition to leadership optimizations, the iterated prisoner’s dilemma model also suggests changes in global climate negotiations. Research by Alessandro Tavoni demonstrates that communication between countries was “vital” and is associated with a sharp increase in rates of disaster avoidance. Although we have facilitated international discussion through structures such as the UNFCCC, America has recently taken steps to close these channels of communication in signalling an intent to withdraw from the Paris Climate Accord. We cannot

---

<sup>15</sup> Page, Edward A. "Distributing the Burdens of Climate Change." Taylor & Francis, 10.

<sup>16</sup> Page, 10.

allow the United States to quietly withdraw into isolationism as it has done so often in the past. Even if the current administration wants to continue to call the Paris Agreement “draconian” and “disastrous,” it is imperative that the United States remain at the table and negotiate a satisfactory alternative in good faith. It does not reflect well upon our nation to renege upon our commitment to protecting the world’s collective future and abandoning the world to suffer its fate.

These international treaties such as the Paris Climate Accord, however, are deeply flawed structurally. Research conducted by Astrid Dannenberg concluded that “the formulation of goals, as long as they are non-binding and costless, is not sufficient to make sure that we stay on the safe side of the threshold [and that] credible and well-balanced commitments [are] much more important.”<sup>17</sup> In this regard, meaningful sacrifices establish genuine commitment are much more effective than the empty and unenforceable promises that countries have typically made. Yet despite each having a vested interest in a successful outcome, we find ourselves locked in a waltz of uncertainty as each nation is paralyzed by the fear that other countries will not reciprocate in kind. Some country needs be the first to step up and demonstrate how highly it prioritizes climate change mitigation. We each don’t want to be the first to commit for fear of being made a fool and taken advantage of, but we are all in this together. It appears as if we have and will continue to reject the premise of our interdependence until a country finally asserts its commitment.

Although most nations are afraid of the consequences of climate change, there exists immense free rider incentives among less developed nations. As stated by Evan Abrams and

---

<sup>17</sup> Dannenberg, Astrid, Andreas Löschel, Gabriele Paolacci, Christiane Reif, and Alessandro Tavoni. "On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds." SpringerLink, 13.

echoed by Peter Wood, smaller, less industrialized nations are motivated to freely emit<sup>18</sup> “since they contribute little to the problem and their efforts to solve it would be negligible.”<sup>19</sup> Yet despite the overwhelming incentive for small nations to free ride, those countries still need to contribute because it discourages major players from defecting. To understand why, we need to look at the ultimatum game, an experiment conducted by various researchers. In this experiment, the first player proposes a way to split an amount of money and the second player can choose to either accept (money is split) or reject (neither receives anything) the proposal, with no chance to make a counteroffer. Rational logic dictates that the second player should accept any proposal greater than zero yet experimental data shows that people typically reject any proposal more uneven than 70/30.<sup>20</sup> Researchers such as Aaron K justify this discrepancy both as an altruistic way to punish greed and exploitation and as a egoistic way to prevent the creation of “a marginalized wealth gap.”<sup>22</sup> Peter Wood argues that “the ultimatum game [shows] that an international environmental agreement is more likely to be stable if it is perceived...to be fair.”<sup>23</sup> This brings us back to the original point: undeveloped countries should contribute in combating global warming in order to more evenly distribute the burden of change and decrease the likelihood of defections by major parties.

That brings up the question of what to do should a party choose to defect. The dominant strategy in the iterated prisoner’s dilemma (tit-for-tat) is successful because it is nice, forgiving, and retaliatory. The tit-for-tat strategy dictates that countries should punish defection by, in turn,

---

<sup>18</sup> Wood, 3.

<sup>19</sup> Abrams.

<sup>20</sup> K, Aaron. "A Brief Refutation of Game Theory." A Brief Refutation of Game Theory.

<sup>21</sup> Wood, 11.

<sup>22</sup> K, Aaron.

<sup>23</sup> Wood, 11.

defecting. However, countries should not directly follow this course of action because “they hurt signatories as much as non-signatories” and “involve impacts that are experienced decades into the future.”<sup>24</sup> Alternatively, Wood proposes that nations link cooperation with trade through a variety of methods such as trade restrictions and border tax adjustments, which were effective in ratifying and succeeding in the Montreal Protocol.<sup>25</sup> These sanctions provide a way to punish defectors without damaging the environment. They make international agreements much more than just empty promises; they are a mechanism that drives the prisoner’s dilemma system towards a cooperative equilibrium.

### **Conclusion**

Without modification, the prisoner’s dilemma is by no means an adequate lense to view climate change mitigation through. To perceiving global warming as a prisoner’s dilemma is to ignore the crucial complexities and nuances that underlie the issue and misconstrue our generation’s most defining issue. These simplifications are problematic because most Americans unknowingly understand climate change mitigation as the prisoner’s dilemma. That means that flaws in the prisoner’s dilemma reflect issues with our perception of climate change. Moreover, studying variants of the prisoner’s dilemma that address those flaws reveals a slough of mistakes that we have made in climate negotiation. By no means are these actions guaranteed to work. But considering how little traction our current strategies have gained, it’s probably worth a try. This time around, we need everyone, from top to bottom, to join in saving our world. In the words of

---

<sup>24</sup> Wood, 14.

<sup>25</sup> Wood, 14.

Shane Koyczan, “this one world is all we get... Forget about the cost. There will be no other thing as worth saving as this.”<sup>26</sup>

---

<sup>26</sup> Koyczan, Shane. ""Shoulders" by Shane Koyczan and The Short Story Long." YouTube.

## References

- Abrams, Evan. "Climate Change Is the Ultimate Prisoners' Dilemma | GRI." Global Risk Insights. May 9, 2013. Accessed April 16, 2018.  
<https://globalriskinsights.com/2013/05/climate-change-is-the-ultimate-prisoners-dilemma/>.
- Althor, Glenn, James E. M. Watson, and Richard A. Fuller. "Global Mismatch between Greenhouse Gas Emissions and the Burden of Climate Change." Nature News. February 05, 2016. Accessed April 16, 2018. <https://www.nature.com/articles/srep20281>.
- Dannenbergh, Astrid, Andreas Löschel, Gabriele Paolacci, Christiane Reif, and Alessandro Tavoni. "On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds." SpringerLink. May 14, 2014. Accessed April 16, 2018.  
<http://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2014/05/ambiguity-ERE.pdf>.
- Diamond, Jared M. *Collapse: How Societies Choose to Fail or Succeed*. New York: Penguin Books, 2011.
- Dyke, James. "Can Game Theory Help Solve the Problem of Climate Change?" The Guardian. April 13, 2016. Accessed April 16, 2018.  
<https://www.theguardian.com/science/blog/2016/apr/13/can-game-theory-help-solve-the-problem-of-climate-change>.
- "Information Provided by Parties to the Convention Relating to the Copenhagen Accord." UNFCCC. Accessed April 25, 2018.  
<https://unfccc.int/process/conferences/pastconferences/copenhagen-climate-change-conference-december-2009/statements-and-resources/information-provided-by-parties-to-the-convention-relating-to-the-copenhagen-accord>.
- Iris, Doruk, Jungmin Lee, and Alessandro Tavoni. "Delegation and Public Pressure in a Threshold Public Goods Game: Theory and Experimental Evidence." Grantham Research Institute on Climate Change and the Environment. March 20, 2015. Accessed April 16, 2018.  
<http://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2015/03/Working-Paper-186-Iris-et-al.pdf>.
- Irwin, Timothy. "Implications for Climate-Change Policy of Research on Cooperation in Social Dilemmas." World Bank Document. July 2009. Accessed April 16, 2018.  
<http://documents.worldbank.org/curated/en/538101468336321961/pdf/WPS5006.pdf>.
- K, Aaron. "A Brief Refutation of Game Theory." A Brief Refutation of Game Theory. August 15, 2010. Accessed April 18, 2018.  
<http://www.sentientdecisionscience.com/ultimatum-coins-of-emotional-fortune-and-a-brief-refutation-of-game-theory/>.

- Koyczan, Shane. "'Shoulders' by Shane Koyczan and The Short Story Long." YouTube. October 08, 2014. Accessed April 25, 2018. [https://www.youtube.com/watch?v=An4a-\\_NjilY](https://www.youtube.com/watch?v=An4a-_NjilY).
- Newkirk II, Vann R. "Is Climate Change a Prisoner's Dilemma or a Stag Hunt?" The Atlantic. April 21, 2016. Accessed April 16, 2018. <https://www.theatlantic.com/notes/2016/04/climate-change-game-theory-models/479340/>.
- Ostrom, Elinor. "A Polycentric Approach For Coping With Climate Change." A Polycentric Approach For Coping With Climate Change: Policy Research Working Papers. October 2009. Accessed April 16, 2018. <https://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5095>.
- Page, Edward A. "Distributing the Burdens of Climate Change." Taylor & Francis. August 8, 2008. Accessed April 16, 2018. <https://rsa.tandfonline.com/doi/pdf/10.1080/09644010802193419>.
- Sagoff, Mark. "Climate Change Is Not a Tragedy of the Commons." Learn Liberty. April 22, 2017. Accessed April 16, 2018. <http://www.learnliberty.org/blog/climate-change-is-not-a-tragedy-of-the-commons/>.
- Sloane, Heidi. "Recycling and a Meat-free Diet Won't Stop Climate Change. What Can You Do?" Medium. March 11, 2016. Accessed April 16, 2018. <https://medium.com/disruptive-design/recycling-and-a-meat-free-diet-won-t-stop-climate-change-what-can-you-do-e0f0980d3c46>.
- Spring, Judith. "Why Climate Change Is Essentially Not a Prisoner's Dilemma." IMSD International Master in Sustainable Development and Corporate Responsibility RSS. November 30, 2015. Accessed April 16, 2018. <http://www.eoi.es/blogs/imsd/why-climate-change-is-essentially-not-a-prisoners-dilemma/>.
- Tavoni, Alessandro. "COP21: Can Game Theory Predict a Global Deal on Climate Change?" Grantham Research Institute on Climate Change and the Environment. December 11, 2015. Accessed April 16, 2018. <http://www.lse.ac.uk/GranthamInstitute/news/cop21-can-game-theory-predict-a-global-deal-on-climate-change/>.
- "UN, United Nations, UN Treaties, Treaties." United Nations. Accessed April 25, 2018. [https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg\\_no=XXVII-7-d&chapter=27&clang=\\_en](https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-d&chapter=27&clang=_en).
- Wood, Peter John. "Climate Change and Game Theory: A Mathematical Survey." Environmental Economics Research Hub Research Reports. July 19, 2011. Accessed April 16, 2018. [https://crawford.anu.edu.au/research\\_units/eehr/pdf/EERH\\_RR62.pdf](https://crawford.anu.edu.au/research_units/eehr/pdf/EERH_RR62.pdf).