Book Proposal for:

Climate Games:
Experiments on the Strategy of Disaster Prevention

Talbot M. Andrews
Department of Political Science
Center for the Study of Democratic Politics
Princeton University
talbota@princeton.edu

Andrew W. Delton
Department of Political Science
Center for Behavioral Political Economy
Stony Brook University
andrew.delton@stonybrook.edu

Reuben Kline
Department of Political Science
Center for Behavioral Political Economy
Stony Brook University
reuben.kline@stonybrook.edu

When your house catches fire, that's a disaster. It's a disaster for you, obviously. But it's also a disaster for the people living around you. In 1666, the Great Fire of London destroyed the homes of nearly 90% of the city's inhabitants. That fire started at a single bakery. The Great Chicago Fire of 1871 began, apocryphally, because Mrs. O'Leary's cow kicked over a lantern. That fire left about 1 in 3 residents homeless and destroyed (in today's dollars) $4.6 billion in property.

So, whether or not you want to do anything about your burning house, your neighbors certainly do because their houses are also on the line. Still, the problem is not simple for your neighbors. If most rush over and put out the fire, a recalcitrant neighbor who stays home gets all the benefits (their house is saved) without paying any costs (the risks of dealing with the fire). This is the classic problem of a *public good*. In a public good, a benefit is produced collectively, but people who did not contribute—like the recalcitrant neighbor—can take a free ride. The logic of free riding is compelling, if cynical: Why pay the costs if you can get the benefits anyway? This is the strategic dilemma of public goods.

The problem of preventing or mitigating climate change disaster is often described as a global public good. The benefits of mitigation efforts are collective and global, so the world as a whole would be better off if all nations reduced their greenhouse gas emissions. Any given nation, however, would be better off if it continued emitting and let other nations take care of the problem. In this way, the prevention of disastrous climate change is a global public good involving all of the world's nearly eight billion inhabitants. This feature alone makes it challenging to solve. We argue, however, that this characterization does not go far enough. Mitigating climate change involves a series of strategic challenges that go well beyond those inherent in a typical public good. **Our proposed book would use economic games to reveal how real people respond to this unique configuration of strategic challenges.**

Economic games present people with carefully controlled, strategic environments that have real stakes (i.e., money) on the line. This makes economic games complementary to other methods, as different methods have their own advantages and limitations. In survey research, for instance, we must take the world as given and see how people respond to it. When answering survey questions, moreover, we cannot know for sure what incentives and problems people perceive themselves or others to be facing. In contrast, using economic games we can create hypothetical worlds that allow us to create and carefully manipulate the problems and real incentives people face.

This allows us to transparently answer a variety of pressing questions: How do all the uncertainties surrounding climate change—the appropriate maximum temperature rise to shoot for, the steps necessary to do so, the economic tradeoffs involved—affect willingness to invest in mitigation? Will people be willing to invest in risky technology, technology with large upsides but also large chances of failure, to prevent disaster? What if such technology has the potential to backfire and cause more problems than it solves?

How will people deal with disasters that they themselves caused? How do they think about the fact that the people deciding how to respond to climate change often are not the people most likely to be affected? How willing are people to make sacrifices to prevent disaster for others?

How will elites and citizens interact to deal with climate change in cases where their incentives diverge? When might citizens fail to trust elites to make the right decisions? When might elites fail to trust citizens?

Our proposed book will examine all of these questions and more. What unifies our book is a commitment to using transparent experiments to understand the social logic of climate change. A few of our studies that we review have been published elsewhere, but primarily we focus on new data. In what may surprise many readers, our results generally reveal an optimistic picture: Players are often quite rational *and* quite self-sacrificing in their strategies for disaster prevention. Nonetheless, because of the reasons we outline in Chapters 1 and 3, climate change is a singularly challenging strategic problem. Though our experimental results reveal a remarkable amount of strategically rational prosocial behavior, substantial hurdles remain.

Chapter 1: The Strategic Complexities of Climate Change

Once your neighbor's house is on fire, the problem and solutions are relatively well defined. The problem is your house might also catch fire. The solution is to douse the flames in water as quickly and thoroughly as possible. Mitigating climate change is nothing like this. For starters, the extent and nature of the problem is difficult to fully articulate. For instance, what is the social cost of current carbon emissions? Carbon, in the form of carbon dioxide (CO_2), is one of the greenhouse gases that contributes most to climate change. Depending on the model, the parameters, and what steps are assumed to have been taken against climate change, estimates for the social cost of carbon range from $22 to $300 per ton of CO_2. How low do we need to keep global mean temperature in order to prevent catastrophic climate change? The Intergovernmental Panel on Climate Change (IPCC) argues that global mean temperature cannot rise more than 1.5°C above pre-industrial levels. On the other hand, an analysis by Nobel Laureate William Nordhaus suggests that the economically optimum rise, factoring in both the benefits of emissions and the costs of mitigation and adaptation, is 3.5°C by 2100 (Nordhaus 2017). Chapter 3 is organized around this theme of uncertainty.

Another strategic complexity is causation. Throughout most of human history, many of the disasters we faced originated outside of our control. Earthquakes, volcanic eruptions, floods or droughts—these could only be reacted to, not stopped before they started. Climate change is different: Human activities are a key driving force behind global warming. Our continued carbon emissions bring us closer to a global tipping point, that is, to a point beyond which the costs of climate change suddenly accelerate. This leads to a tension: We emit CO_2 because of the benefits of burning fossil fuels. Energy allows greater production, which allows greater standards of living. If history stopped now, the typical human would certainly have been better off from extensive CO_2 emissions. The problem is the more we emit, the greater the future problem becomes. Chapter 4 is organized around this theme of self-causation and the tension between the benefits of economic development and the costs of the emissions arising as a by-product of development.

A further complexity is that the primary actors deciding how to mitigate climate change are not necessarily the people who will be most affected by climate change. There is a multi-generational component: We in the present gain many benefits from emissions; it will be our descendants who will pay the lion's share of the potential costs should we fail to prevent the problem. There is also a geographical component: It will be politicians in rich nations who make many of the most consequential decisions, even though it is people in the developing world who stand to be most affected. Generally, people with less at stake are making mitigation decisions for others. Chapter 5 is organized around this theme of deciding for others.

The final complexity we consider is that decision-makers often have different incentives, not just weaker ones, than the people they are making decisions for. Elite actors are likely better informed about the nature of the problem and how the problem interacts with government institutions. For instance, elite actors are likely to have read the 32 pages of the IPCC 2014 summary report for policy makers and perhaps the 167 pages of the full 2014 synthesis report; most people will not have done so. This has at least two relevant consequences. First, leaders might manipulate the information or the choices they make for their own benefit. Second, even if a leader has incentives that fully align with citizens, they may need to behave in paradoxical ways in attempts to anticipate and adjust for citizens' potentially incorrect beliefs. Elites' incentives or information may not align with citizens'. Chapter 6 is organized around this theme.

These four issues—massive uncertainty, a self-created problem, decision-making for others, divergence between elites and citizens—are not individually unique in climate change. But taken together (and with other complexities we do not address in detail) they make climate change stand out as special problem. Moreover, these complexities interact. Although we have separated them for convenience into distinct chapters, themes that play a major role in one chapter will crop up as the supporting cast in other chapters. Before we begin detailing the various strategic challenges that climate change represents, we first discuss why an experimental economic approach is particularly well-suited to this challenging context. This is the theme we turn to next.

Chapter 2: Creating Hypothetical Worlds

We examine the strategy of disaster prevention using a series of economic games. Economic games allow us to create transparent, strategic environments in the lab. They do this by presenting players with a clear set of rules for how their decisions, individually and collectively, determine how much they earn in real money. Sometimes the rules and payoffs are designed to mimic real-world strategic problems; other times they are designed to create strategic environments that do not exist, at least yet, in the real world. Our work uses experimental economic games to explore both types of worlds.

One economic game that many researchers are already familiar with is the *public goods game*. This game is designed to mimic situations, like the collective fire prevention discussed previously, where everyone is better off if everyone cooperates, but for any given person it is individually rational to free ride and let everyone else do the work. Typically, players in laboratory public goods games are each given a personal sum of money, perhaps $10. Players choose whether to invest their money in a group account that multiplies the original contribution and thereby generates a surplus for the group, or to simply keep their money for themselves. Full contributions are better for everyone, compared to no one contributing. But the game is set up such that each player is individually better off if they free ride and hope others contribute.

When researchers use the public goods game to model climate change mitigation, however, they use a special version called a *threshold* public goods game. In a threshold game, as the name implies, the players are given a threshold which is a monetary amount. Each player is given a sum of money they can keep for themselves or contribute to their group's threshold. Players get rewarded if and only if they contribute enough as a group to meet their threshold. For instance, a group of 4 players, each with $10, might face a threshold of $20. If they collectively contribute $20, each player will earn a bonus of $10 (in addition to keeping what they did not contribute). If they fail to meet the threshold, they do not get any additional bonus *and* they forfeit the money they contributed to the threshold. Notice that any given player would prefer that all the other players share the burden of meeting the threshold. A player who keeps all their personal money when the group otherwise meets the threshold would earn $20, whereas the other members of their group would earn on average $13.33. This illustrates the tension between individual and collective interest in public goods games. A real-world threshold example is academic research: Only if a research team successfully brings a project across the "threshold" of publication do they gain substantial benefits from all of their work. (Unlike a typical public goods game where it is *always* selfishly rational to free-ride, a threshold game has different game theoretic properties. Importantly, depending on the other players' behavior, it *can* be individually rational for a given player to contribute.)

To more closely model climate change, researchers use a variant of the threshold game. In this version, meeting the threshold does not earn players a bonus, but instead prevents a preexisting disaster from occurring. **Throughout the book we will call this the *disaster game*. This is the central game around which all our experiments are built.** In a typical disaster game, each player again starts with a pot of money. When the game begins there is a 90% chance that players will lose all of their money. But players also receive a climate threshold, a monetary amount. If, as a group, they contribute enough to meet the threshold, everyone gets to keep whatever money they have left—disaster is averted.

The structure of this game is designed to match the way many climate scientists think about climate disaster. The costs of climate change do not increase smoothly as the global temperature increases. Instead, there is a relatively crisp point—a tipping point—at which climate change related disasters become much worse. The game simulates this with its all-or-none approach to disaster. Admittedly, the real problem is likely not quite so sharp. But this illustrates a necessary feature of games: They must be sufficiently simplified so that real people can understand and play them. Nonetheless, even with simplification, a well-designed game will capture the key strategic problem of interest and isolate its causal effect on behavior.

Why turn to experimental games? Why not simply run surveys asking people how they think about climate change and how to stop it? We do not think this is an either/or question. Both approaches are useful for understanding how people think about climate change (in fact, our team has conducted surveys on this topic). In our view, the value-added from economic games is that we can precisely control and create the strategic environment that players face. In the real world, some people might perceive the problem one way, some might perceive it another; even experts disagree about details. Mapping this is, of course, an important endeavor and survey research is well positioned for it.

Economic games, however, allow us to do something different. In games we can know exactly what problems and material incentives players face, as can the players. Thus, if players react one way in one experimental condition and a different way in another experimental condition, we know what feature of the strategic environment led to that difference—because we created it. Moreover, because we can quantify the material incentives the players face, we can use game theory to predict how a hypothetical rational and payoff-maximizing player would behave. We can use this as benchmark against which to compare real player behavior.

Also important is that we can create strategic situations in the lab that have no analogue in the real world, at least not yet. For instance, although there is a lot of talk of large-scale geoengineering, no serious projects are days away from being ready to launch—even if we wanted to, we could not enact a program of geoengineering tomorrow. In the lab, however, we can present players with a game that simulates the possibility of geoengineering and its strategic implications. In this way, we can get a sense of how real people would approach this strategic problem, even if the problem has not yet arrived in the real world. This is a key advantage of the experimental economics approach (Morton and Williams 2008).

Finally, games focus the mind and serve as a thought experiment for the reader: Given the rules and the stakes, what would you do in these games?

Chapter 3: Uncertainty All the Way Down

Anthropogenic climate change is at this point scientifically undeniable. Still, the same science recognizes that humankind's impact on the climate is a product of a highly complex political, economic, and environmental system, with serious uncertainty about the timing, extent, and distribution of the effects of anthropogenic climate change. Even if we could (somehow) perfectly predict the social part of the system, that is, what people throughout the world would do, serious uncertainty would remain about the geo-physical effects of the continued accumulation of greenhouse gas emissions. **In other words, the real-world analogues of the features of the disaster game—the size of the threshold, the risk of disaster, and so on—are highly uncertain.**

A great example is provided by the experiments of Manfred Milinski and colleagues (2008). This work kicked off the use of the disaster game as a window into the strategy of climate change. In their version of the disaster game, players faced the possibility that disaster would wipe out all their money. But, if players on average contributed half of their personal money to the threshold, disaster would be averted. Milinski and colleagues manipulated the probability of disaster. Depending on the condition there was a 90%, 50%, or 10% chance that everyone in the group would lose their money if the threshold was not met. These researchers find that players contribute less when the probability of disaster is lower: Whereas players facing a 90% probability tend to contribute about enough to meet the threshold, when facing a 10% probability they contribute only about 60% of what is needed. Ultimately, half the groups facing a 90% probability successfully protected the (simulated) environment, whereas none of the other groups facing a 10% probability did. In another example, Dannenberg and colleagues (2015) manipulated whether players knew the exact value of the threshold (as is typical in the disaster game) or whether there is uncertainty about its true size; they find that groups are much less successful at protecting the environment when there is uncertainty in the cost of meeting the threshold (see also Barrett and Dannenberg 2012).

This literature demonstrates the appeal of experimental economic games to better understand the uncertainty that pervades the problem of climate change. Although games cannot of course be directly extrapolated to global diplomacy, Milinski and colleagues' results suggest that if citizens perceive that the probability of disaster is low, they may not support significant mitigation efforts. Dannenberg and colleagues show that there is tension between scientific accuracy (i.e., communicating the real uncertainties) and citizen action.

The above research focused on uncertainty surrounding the *problem*. Our own research discussed in this chapter focuses on another type of uncertainty: uncertainty over potential *solutions*. Once again, the "world building" aspect that the game theory laboratory offers allows us to test behavioral responses to technologies, institutions, and polices that do not yet exist.

What kind of technology are people willing to invest in to mitigate climate change? For instance, if risky technologies—ones with a high upside but also a high chance of failure—are required, will people invest in them? Indeed, according to the IPCC 2014 synthesis report, many climate change mitigation scenarios rely on high-risk technologies, such as carbon capture and storage, that have never been successful at a wide-scale.

We created a version of the disaster game designed to test under what circumstances people would be willing to invest in risky technology. In this version of the game, players received two pots of money. One was relatively large and represented things like infrastructure, which can be damaged by climate change but cannot be meaningfully used to stop it. The other pot was a "personal account" of liquid funds that could be spent to stop climate change.

Within a group of four players, each player had to decide whether to give their entire personal account to the group or whether to keep their entire personal account for themselves. If they contributed, they had to decide what kind of (simulated) technology to invest in. They could make a certain investment (representing investment in incremental tech like solar or wind power) or a risky investment (representing investment in high-risk/high-reward tech like carbon capture and storage). A certain investment merely contributed the personal account to the group's threshold. But a risky investment was a gamble: There was a 50% chance it would double before going toward the group's threshold and a 50% chance it would disappear.

How would you invest? If you're like most people, your choice would likely depend on the threshold your group faces. If the problem is easy and the threshold is inexpensive, then your best bet is to make a certain contribution. But if the problem is hard and the threshold is especially high, only with at least some players making risky contributions is it even possible to prevent disaster. For instance, if each player has a $4 personal account, the maximum the group can contribute through direct contributions is $16. But if the threshold requires $24, direct contributions won't cut it. In fact, in this example, the best thing for all players to do is invest in risky technology.

Across four studies, some with students in our lab and others with diverse online samples, we found that players were quite good at making useful decisions (Andrews, Delton, and Kline 2018, *Nature Climate Change*). Participants did indeed tend to favor riskier investment when only risky investments could solve the problem. This suggests that citizens might be willing to support investment in risky technology, to the extent that they believe such technology is needed to ameliorate climate change.

The political, economic, and environmental problems posed by climate change are full of uncertainty. The collection of studies in Chapter 3 highlight that uncertainty can make it more difficult for people to coordinate around a climate change solution. Optimistically, however, we also find that people can recognize when taking risks for the climate might be warranted.

Chapter 4: Flirtin' with (Self-Created) Disaster

An isolated natural disaster, such as a volcanic eruption, occurs regardless of human choice. A human conflict, like war, is purposefully destructive. **Climate disaster, however, is a consequence of our own actions, but an unintended one.** In our still de-carbonizing global economy, the size of the threshold, the likelihood of climate disaster, and the resources available to prevent disaster are all determined by humans through our economic development activity. In other words, climate disaster is a self-created disaster.

Moreover, these unintended consequences are harmful not only to ourselves, but also create negative externalities, that is, consequences for others who had no say in creating them. This includes rich countries creating negative externalities for contemporaries in developing countries and the present generation creating negative externalities for future generations by continuing to emit CO_2 at the current rate (for a game version, see Jacquet et al. 2013). At the same time, economic growth per se confers *positive* externalities on future generations in terms of increased living standards. This tradeoff between wealth creation and the increasing likelihood of disaster permeates the strategic landscape of climate change. In this chapter we discuss several studies that more closely examine the implications of the self-created nature of the climate change problem.

One of us (Kline), along with our colleagues Nicholas Seltzer, Evgeniya Lukinova, and Autumn Bynum conducted a series of studies using a variant of the disaster game in which the players' initial endowments, the climate threshold, and the probability of disaster are all self-created through the players' own choices in a preceding economic development phase (Kline et al. 2018, *Nature Human Behaviour*). The more wealth the players generate in the economic development phase, the richer they are but the higher the climate threshold and the greater the disaster probability in the subsequent disaster game. When these self-created disasters are compared to identically parameterized disaster games in which there is no preceding economic development phase, the probability that the group succeeds falls by nearly half. In other words, a problem that is self-created leads to worse outcomes than an identical problem that was simply handed to players.

Due to the self-created nature of the problem, those who have greater wealth are also more responsible for the severity of the climate disaster. Because of the careful control afforded by an experimental approach, it is transparent to all which players are most responsible for climate problems: However much wealth a player earned in the economic development phase, they increased the climate threshold by an amount equal to 53% of their accumulated wealth. Still, despite this determinism, there is room for reasonable disagreement about how to equitably distribute the burden of meeting the threshold. At the very least it represents a factor that complicates decision-making. This intuition parallels a key equity principle in international climate negotiations known as *common but differentiated responsibilities*. The repeated failure of international negotiations to arrive at an effective and enforceable solution to the problem is often attributed to disagreements over precisely how to differentiate responsibilities for mitigation.

In our game, we observed that disagreement and reduced cooperation resulted even when opportunities for economic development were transparently equal and open to all; this was the baseline version described above. We also ran a version where we introduced asymmetric opportunities for economic development (see also Tavoni et al. 2011 on inequality in this game). Here, we randomly allowed some players—who were called *early developers*—an economic development period twice as long as players assigned to a *late developer* role. This setup meant that the maximum amount of wealth that could be accumulated by late developers was only half as much of that as early developers. (In the baseline version, all players were functionally "early developers.") This mimics the effect of historically uneven economic development. This change eroded cooperation and success even further. This was true even though early developers showed some prudence: They accumulated less wealth than players in the baseline condition and thereby contributed less to the problem (reducing their wealth accumulation by 18%) *and* they contributed a significantly larger proportion of their wealth toward the threshold (56%, compared to 52% in the baseline condition; the contribution that just offsets the negative effects of one's own wealth accumulation is the 53% mentioned above). Nonetheless, the prudence of the early developers was offset by the choices made by late developers. Compared to the baseline, late developers worsened the problem by increasing wealth accumulation by 10%, and decreasing contributions by 18%. Merely introducing inequality of opportunity reduces cooperation in the self-created disaster game, as it provides yet another dimension of potential disagreement over burden-sharing obligations.

The above results came from student players in the United States. But research has shown people tend to choose distributional equity principles that would benefit their own country. Would a similar sample of students from a country with a different historical emissions trajectory than the US behave differently? We conducted an identical experiment with Chinese undergraduate students and the results were remarkably similar. The primary difference was that Chinese players, across all conditions, appropriated higher levels of wealth than their American university counterparts (an average of 24% more across all conditions). Still, just like the Americans, Chinese students behaved differently depending on their development status: The early developers generated less wealth and contributed proportionally more compared to the baseline, but the late developers appropriated more and contributed less. As before, the behavior of late developers offset the prudence of early developers and led to worse outcomes here than in the baseline condition.

We have so far seen the challenges of the human-created nature of climate disaster in cases where players will be *personally* affected by the disaster they create. Climate change, however, is a social dilemma of *global* proportions. As a result, **individuals are also routinely making decisions that will, in effect, create the parameters of the disaster game for *others*.**

With our colleagues Alessandro Del Ponte and Nick Selzter, we ran a game designed to simulate creating a problem that will be faced by others (Del Ponte et al. 2017, *Journal of Politics*). The key difference here is that instead of determining the parameters of the disaster game for one's own group, players in some conditions are creating the parameters of a disaster game for another group to play. We also varied the identity of the other groups—in one condition those other groups are based in the United States and in another the other groups are based in India. (The focal players who could pass on problems were always in the US.) When groups created potential climate disasters for others, they tended to be less restrained in their accumulation of wealth than when creating them for themselves. However, this effect was not overwhelming: Players accumulated only about 15% more wealth when doing so creates problems for others rather than for themselves. Perhaps surprisingly, the country location of these groups did not affect behavior. In other words, players in the United States were neither more nor less selfish when their negative externalities were borne by other groups based in the United States than when these other groups were based in India.

This chapter shows another advantage of using the disaster game approach: Examining the ethics involved in climate change. Experimental economic games have long been used to investigate preferences for the overall distribution and allocation of resources. Experimental control makes the mappings from collective actions to individual payoffs stark and transparent, allowing for sharp measurement of preferences for fairness and equity. While in this chapter we have investigated people's ethical decisions involved in *creating* disaster for others, in the next chapter we use a disaster game approach to investigate how people might *prevent* disasters faced by others.

Chapter 5: Preventing Disaster for Others

The tiny Pacific Island nation of Kiribati might disappear from the map by 2100 due to rising waters. In the meantime, its population is likely to face damaged reefs, storm surges, increased erosion, and less productive fisheries. Starting in 2003, the World Bank helped finance a $17.7 million adaptation program, which included building seawalls and planting mangrove trees. But a later study in 2014 found that at least some of these efforts were counterproductive and hastened erosion (Ives 2016).

Kiribati illustrates **the profound ways in which climate change decisions are often made by one set of people and affect others**. Kiribati has a population of about 110,000 (CIA World Factbook 2019). If their nation disappears due to human-caused climate change, it can hardly be the case that this tiny population was in any meaningful sense responsible; their share of CO_2—past, present, and future—is essentially nil. Whether disaster will be averted for them depends entirely on the choices of other countries.

Kiribati also illustrates **the possibility that steps taken by outside parties could backfire and make the problem worse**. The lead author of the study critiquing the adaptation efforts remarked, "The idea that an outside organization can just come in with money, expertise and ideas and implement something easily is naïve."

In this chapter, we turn these problems into economic games. First, we examine whether people are willing and able to make decisions about disasters for others. We present players with a situation in which there is an objectively better choice they could make on behalf of others. Our question is whether or not they will pay the costs to determine and enact the correct choice.

A long tradition in political science suggests that players will not. This classic view originated in discussions of voting (Downs 1957). Consider that in a large democracy like the United States, millions of votes will be cast in an election. The chance that any one vote will be decisive is nearly zero. But getting informed about the election and going out to vote has costs (e.g., the time it takes to travel). What should you do if you are primarily concerned with your own personal welfare? As your vote has a vanishingly small probability of being decisive, there is no point to paying these costs; you might as well stay home. This argument also implies that because there are no expected material gains to voting, there is also no point to taking the time and energy to become informed about the candidates and issues. Why become informed if doing so will not translate into action?

What's striking about this argument is that it applies even if the voter would be dramatically and personally affected by the election outcome. It's easy to see, then, that the problem would be magnified if the decision has to be made on behalf of others. If it isn't worth paying the costs to get informed about policies that affect you, then why would you pay the costs to get informed about policies that only affect others? When a decision primarily affects other people, this classic view would predict that people will not be motivated to make the decision nor to become informed about what choice is best.

Despite what this theory predicts, people do go to the polls (plus donate to political groups, go to protests, and so on). One explanation for this that people have *social preferences*, a willingness to trade off personal welfare to benefit others (a variety of models of social preferences have been formalized in economics and political science). Surprisingly, although there have been decades of research on social preferences, this work has primarily focused on the basic question of *whether* people will be generous. Our work moves beyond this by asking *what form* generosity will take: Will people in their generosity also be able to make a useful decision? This is not as obvious as it seems: Research on charitable giving shows that most people do no research on whether a charity they plan to donate to is effective at its stated goals, and are thus seemingly motivated by the "warm glow" of altruism, rather than a more consequentialist view of their potential impact (Andreoni 1995).

To study this, we use the game described in Chapter 3: Players must decide whether to make a certain contribution which will directly go to the threshold, or a risky contribution which has an even chance of being doubled or being lost entirely. As in that study, the right choice depends on the threshold players are facing. The catch in the studies for this chapter is that players' choices do not affect themselves; instead, they affect another group. In this version of the game, players have nothing material to gain by giving, nor do they receive any benefits by thinking through what type of giving would be best. Instead, they have to decide whether and how to help another group.

In our first experiment we find that the overwhelming majority of players do choose to contribute to help the other group (~80%). But this merely replicates past work. Will their generosity be useful? Yes, players were more likely to make risky contributions for others when riskiness was necessary and certain contributions when it was not. In a second study, we showed that players are willing to buy information that tells them what the most useful choice is, and many change their decision based on this information. Players are willing to be generous and to do so in a useful way. Of course, this study provides a best-case scenario. It is reasonably clear in this game how to help others. What happens when how best to help is less clear, when there is the possibility that helping might actually backfire?

In this next study, players again have two pots of money, a large endowment that can be damaged by disaster and a smaller personal account that they can spend to stop disaster. The twist in this study is that one player is a "policy maker" and can decide whether to enact simulated geoengineering. Scientists and politicians debate the merits of real geoengineering—such as large-scale seeding of the atmosphere with chemicals—because although the benefits might be huge, the climate system is complex and hard to predict, and such efforts might backfire. In our game, when the policy maker chooses to use geoengineering, it can either succeed or make everyone worse off. If it succeeds, disaster is averted and the group can keep all its money without needing to contribute to the threshold. But if geoengineering backfires, the group pays a separate and large penalty, *and* they still face the same threat of disaster as before. The policy maker responds to a series of possible chances that geoengineering succeeds (e.g., 90% chance of success/10% chance of backfiring, 50% success/50% backfire, etc.). Across these probabilities, there are some points where it is optimal for the group if the policy maker uses geoengineering and some where it is optimal not to.

Using this experiment, we can see whether people are too cautious about (simulated) geoengineering or too eager. Between-groups, we manipulate whether the policy maker is a third-party or is a member of the group. If the policy maker is a member of the group, then they face the consequences of their choice; as a third-party, the policy maker is completely unaffected by what they decide. Thus, compared to people who are directly affected, we can observe whether third-parties are more or less circumspect about using geoengineering.

Chapter 6: Can Elites and Citizens Trust Each Other?

"Donald Trump has tweeted climate change skepticism 115 times." So begins a 2017 Vox article (Matthews 2017). A 2013 tweet read, "We should be focused on magnificently clean and healthy air and not distracted by the expensive hoax that is global warming!" One in 2014 read, "Any and all weather events are used by the GLOBAL WARMING HOAXSTERS to justify higher taxes to save our planet! They don't believe it $$$$!" Tweeting during his administration, in 2018 he opined, "Brutal and Extended Cold Blast could shatter ALL RECORDS – Whatever happened to Global Warming?"

Although Trump is himself an elite, his rhetoric channels **a worry that some citizens have: Concern over climate change is exaggerated or outright fabricated because scientists, politicians, and other elites benefit from doing so.** Politicians get more tax dollars to control; scientists gain in appointments, prestige, and grant money.

In fact, this assumption of citizen suspicion is built into research on designing economic mechanisms to combat climate change. For instance, in experimental research on (simulated) cap-and-trade systems, researchers assume that citizens will be suspicious of mechanisms where the government generates a profit (e.g., Franciosi et al. 1993; Ledyard and Szakaly-Moore 1994). Indeed, a rational citizen should recognize this incentive and might refuse to participate (Miller and Hammond 1994). Altogether, this suggests that citizens should be sensitive to institutional mechanisms whereby elites capture benefits for their own ends.

On the other hand, research in political science argues that citizens are often uninformed (Delli Carpini and Keeter 1993) and process information in a way biased by their partisanship rather than for accuracy (Jerit and Barabas 2012). Moreover, voters favor short-term economic gains over more efficient long-term risk prevention, rewarding disaster relief spending but punishing spending on prevention (Healy and Malhotra 2009). This generally suggests citizens have a poor grasp on the political economics of disaster response. Voters might not care or be able to recognize when elites can enrich themselves based on institutional design.

Here, we create an economic game to test whether citizens are sensitive to the design of institutions that may or may not enable elites to enrich themselves at the expense of the citizens. Will citizens disbelieve elites who have a stake in inefficiency and therefore contribute less to climate change mitigation? Answering this question is important for real-world mitigation efforts. If citizens are sensitive to institutional differences *and* they come to believe that institutions favor elites (even if such beliefs are false), they will not support climate change mitigation efforts.

In this game, players face a 100% probability of disaster if they fail to meet their group's threshold. Unfortunately for the players, they do not know exactly what the threshold is, though they know it will be drawn from one of five possibilities. A fifth player, the elite, does know the exact value of the threshold and can communicate this to the other players. (The elite cannot personally contribute to the threshold.) However, the elite is not required to be truthful; they can communicate any of the five possible values.

In the control condition, the elite faces the same problem as the other players: Everyone, including the elite, has a pot of money that will be lost if disaster occurs. In this condition, all players' incentives are aligned. Rational elites should tell the truth and players should believe the elite and therefore contribute sufficiently to meet the threshold.

In the inefficiency condition, the incentives are not aligned. The elite still has a pot of money they stand to lose if their group fails to prevent the disaster. However, if the other players contribute more than was strictly necessary to meet the threshold, the elite gets to keep a portion of the excess contributions. Thus, elites have an incentive to exaggerate the threshold that the group faces. Rational, payoff-maximizing elites should always tell their group they face the highest threshold and the other players should disbelieve them and should contribute as if they received no information from the elite. (We set the game up so that in this case the other players should act as if they face the middle of the five thresholds.)

Of course, players may not be rational, or they may have social preferences. What do real players do? Real players assigned to the elite role in this game show a mix of behavior. On average, elites in the inefficiency condition exaggerated the size of the threshold by nearly 30%. However, even in the inefficiency condition, 52% of elites reported the actual threshold. If the other players appropriately anticipate the general behavior of elites in this game, they should be somewhat skeptical, but not unduly so.

In the inefficiency condition, the other players are more skeptical of the elites than in the control condition—especially when the elite claims it is very expensive to meet their threshold. When an elite in the inefficiency condition says that the cost of meeting the threshold is the most expensive possible value, people think the elite is exaggerating the cost by approximately 30%. They accurately anticipate the elite's stake in inefficiency. However, when elites in the inefficiency condition say the threshold is actually *in*expensive, 63% of players believe the elite is telling the truth. Unfortunately, the overall effect is that the possibility that elites can enrich themselves undermines successful disaster prevention. Groups in the inefficiency condition were less likely to contribute enough to stop the disaster than in the control condition.

This study examined a case designed to model citizens distrusting elites. But **can elites sometimes fail to trust citizens? This is a major issue in discussions of "moral hazard" in efforts to mitigate climate change.** Moral hazard is a sort of ironic effect whereby when the consequences of some risk are mitigated, people become more likely to take that risk. For instance, a person who gets health insurance may become more likely to engage in dangerous extreme sports like mountain climbing. The problem of moral hazard here would be if people become unduly risk-seeking after becoming insured, such that they would have been better off *not* getting insurance and *not* going climbing.

In the context of climate change, one source of worry about geoengineering is that it creates moral hazard in citizens. If citizens think that moon-shot geoengineering efforts will fix the problem, there is little incentive for them to support more piecemeal (but still necessary) mitigation efforts. An extreme version of this worry would be that merely raising geoengineering as a hypothetical future possibility could make things worse. One reason that geoengineering could create a problem is that citizens may be overoptimistic compared to elites. Elites, such as scientists and policy makers, are likelier to have a nuanced understanding of the risks, uncertainties, and pitfalls associated with geoengineering efforts. Citizens, however, may place too much faith in their effectiveness.

We have designed a game to simulate this dynamic. An elite player has the option to implement geoengineering. If geoengineering succeeds, the disaster is resolved; if it fails, the game continues on as before. (To be clear, unlike the previous geoengineering study, here there is no penalty if geoengineering fails. If it fails, it's merely ineffectual rather than detrimental.) The remaining players, the citizens, will have the opportunity to contribute to a threshold to prevent disaster. If disaster occurs, it wipes out the money of *both* the elite and the citizens.

The key to this game is that players have differential access to information. Elites know the true probability that geoengineering will be successful. The citizens do not. Instead, the citizens get a noisy signal about the likelihood of success, one that might be nearly accurate or that might be overly optimistic. Moreover, the citizens do not know whether geoengineering actually succeeded or failed; they only know whether the elite chose to use it and the (noisy) probability it will succeed. The elites know both the true probability geoengineering will be successful, and whether or not the citizens have accurate or overly optimistic information about geoengineering's potential success.

In principle, elites should always use geoengineering. In this game it has only upsides and no downsides. However, because the citizens' beliefs might be optimistically biased, elites might paradoxically withhold geoengineering. If the citizens assume geoengineering will be successful—due to their inaccurate, optimistic beliefs—they may fail to contribute sufficiently and therefore fail to prevent disaster. An elite who sees this coming might decide it's better to step out of the way and do nothing, rather than give the citizens false confidence that the problem is solved.

Chapter 7: A Strategic Hope

Dealing with climate change is complex and one general concern is that people are unwilling and unable to overcome these complexities. Despite this conventional wisdom, we believe our results warrant a cautious optimism. Our players seemed able to make good decisions and to strike a balance between their own material benefits and others'. For instance, in our game that allows players to make certain or risky contributions (Chapter 3), players tend to choose the most appropriate type of contribution. Moreover, they make these useful decisions even when doing so is personally costly and only helps others (Chapter 5).

Although players did tend to create worse problems when the problems were passed along to others, this difference was fairly modest (Chapter 4). And the problems were no worse when players passed to players in another country half a world away compared to passing to others in their home country. It's also true that elites behaved badly, and other players expected this, at least when the institution enabled elites to exploit average citizens (Chapter 6). But again, the bad behavior was relatively modest. Many elites were honest regardless of their incentives to be corrupt, and many players assumed elites would be honest even under bad institutions.

Throughout, we've used economic games to simulate strategic problems humanity faces now and ones humanity will soon face. Our results show how real people respond to these problems when real money is on the line. Beyond that, we believe thinking through these stark and simple games focuses the mind on the essence of the problems. We hope that readers—scientists, policy makers, and perhaps even politicians—will come away with a deeper understanding having considered for themselves how they would play the game.

References

Andreoni, James. 1995. "Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments." *The Quarterly Journal of Economics* 110(1): 1–21.

Andrews, Talbot M., Andrew W. Delton, and Reuben Kline. 2018. "High-Risk High-Reward Investments to Mitigate Climate Change." *Nature Climate Change* 8: 890–94.

Barrett, Scott, and Astrid Dannenberg. 2012. "Climate Negotiations under Scientific Uncertainty." *Proceedings of the National Academy of Sciences of the United States of America* 109(43): 17372–76.

Dannenberg, Astrid et al. 2015. "On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds." *Environmental and Resource Economics* 61(3): 365–83.

Del-Ponte, Alessandro, Andrew W. Delton, Reuben Kline, and Nicholas A. Seltzer. 2017. "Passing It Along: Experiments on Creating the Negative Externalities of Climate Change." *The Journal of Politics* 79(4): 1444–48.

Delli Carpini, Michael X., and Scott Keeter. 1993. "Measuring Political Knowledge: Putting First Things First." *American Journal of Political Science* 37(4): 1179.

Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy* 65(2): 135–50.

Factbook, CIA World. 2019. *Kiribati*. Washington, DC.

Franciosi, Robert, R. Mark Isaac, David E. Pingry, and Stanley S. Reynolds. 1993. "An Experimental Investigation of the Hahn-Noll Revenue Neutral Auction for Emissions Licenses." *Journal of Environmental Economics and Management* 24: 1–24.

Healy, Andrew, and Neil Malhotra. 2009. "Myopic Voters and Natural Disaster Policy." *American Political Science Review* 103(03): 387–406.

Ives, Mike. 2016. "A Remote Pacific Nation, Threatened by Rising Seas." *The New York Times*.

Jacquet, Jennifer et al. 2013. "Intra-and Intergenerational Discounting in the Climate Game." *Nature Climate Change* 3(12): 1025.

Jerit, Jennifer, and Jason Barabas. 2012. "Partisan Perceptual Bias and the Information Environment." *The Journal of Politics* 74(3).

Kline, Reuben, Nicholas Seltzer, Evgeniya Lukinova, and Autumn Bynum. 2018. "Differentiated Responsibilities and Prosocial Behaviour in Climate Change Mitigation." *Nature Human Behaviour* 2(9): 653–61.

Ledyard, John O., and Kristin Szakaly-Moore. 1994. "Designing Organizations for Trading Pollution Rights." *Journal of Economic Behavior & Organization* 25(2): 167–96.

Matthews, Dylan. 2017. "Donald Trump Has Tweeted Climate Change Skepticism 115 Times. Here's All of It." *Vox*.

Milinski, Manfred et al. 2008. "The Collective-Risk Social Dilemma and the Prevention of Simulated Dangerous Climate Change." *Proceedings of the National Academy of Sciences of the United States of America* 105(7): 2291–94.

Miller, Gary, and Thomas Hammond. 1994. "Why Politics Is More Fundamental Than Economics." *Journal of Theoretical Politics* 6(1): 5–26.

Morton, Rebecca B., and Kenneth C. Williams. 2008. "Experimentation in Political Science." In *The Oxford Handbook of Political Methodology*, eds. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier. , 339–56.

Nordhaus, William D. 2017. "Projections and Uncertainties about Climate Change in an Era of

Minimal Climate Policies." *NBER Working Paper*: 22933.

Tavoni, Alessandro, Astrid Dannenberg, Giorgos Kallis, and Andreas Löschel. 2011. "Inequality, Communication, and the Avoidance of Disastrous Climate Change in a Public Goods Game." *Proceedings of the National Academy of Sciences of the United States of America* 108(29): 11825–29.