Collective Risk and Distributional Equity in Climate Change Bargaining

Aseem Mahajan[*], Reuben Kline[†], Dustin Tingley[‡]

Draft as of January 20, 2020

Abstract

International climate negotiations occur against the backdrop of increasing collective risk: the likelihood of catastrophic economic loss due to climate change will continue to increase unless and until global mitigation efforts are sufficient to prevent it. We introduce a novel alternating-offers bargaining model that incorporates this characteristic feature of climate change. We test the model using an incentivized experiment. We manipulate two important distributional equity principles: capacity to pay for mitigation of climate change and vulnerability to its potentially catastrophic effects. Our results show that less vulnerable parties do not exploit the greater vulnerability of their bargaining partners. They are, rather, more generous. Conversely, parties with greater capacity are less generous in their offers. Both collective risk itself and its importance in light of the recent Intergovernmental Panel on Climate Change report make it all the more urgent to better understand this crucial strategic feature of climate change bargaining.

Keywords: climate change, collective risk, equity, laboratory experiment, bargaining

[*]Department of Government, Harvard University, mahajan@g.harvard.edu

[†]Department of Political Science & Center for Behavioral Political Economy, Stony Brook University, reuben.kline@stonybrook.edu

[‡]Department of Government, Harvard University, dtingley@gov.harvard.edu

1 Introduction

A recent report by the Intergovernmental Panel on Climate Change (IPCC) underscores the urgent need to address climate change, noting that the window of time over which countries must initiate climate policies to avoid its worst consequences is as small as 12 years. Given the current stock of greenhouse gas (GHG) emissions, countries must now undertake drastic efforts to keep warming below the recommended 1.5°C. Delaying action any further will increase countries' risk of catastrophic economic consequences of climate change (IPCC 2018).

We investigate two questions related to the scenario laid out by the IPCC. How does an increase in the risk of economic catastrophe affect individuals' willingness to bear the costs of climate change mitigation? And, in light of global inequality in countries' vulnerability to climate change and their capacity to prevent it, how do differences in these two factors moderate the extent to which individual preferences respond to increased economic risks?

The consensus view of policymakers and scholars is that reducing the catastrophic effects of climate change requires an effective and sustainable international agreement to collectively reduce global emissions. Therefore the imminent threat of catastrophic climate change lends urgency to understanding how countries' increasing and differential vulnerability to collective risk affects the costs they are willing to accept to prevent it. Strategically, collective risk impacts global efforts to mitigate climate change, and is thus a critical feature of international climate negotiations. Despite its importance, the topic remains under-studied. In this study, we introduce a flexible bargaining framework that incorporates increasing collective risk into an otherwise standard alternating-offers bargaining model.

Recent estimates of the value at risk from unmitigated climate change suggest expected costs of $2.5 trillion, with substantial risk in the tail (Dietz et al. 2016; Weitzman 2011). As a stylized representation of these risks, Milinski, Sommerfeld, et al. (2008) introduce the concept of "collective risk" to describe the a threat of widespread and catastrophic economic loss posed by unabated climate change (Alley et al. 2003; Schellnhuber 2006). Given current trends, the likelihood of widespread, collective risk will continue to increase unless countries sufficiently reduce net global GHG emissions. Of course, not all countries, regions, and individuals are equally vulnerable to these risks. Likewise, the resources to prevent climate change are also unevenly distributed.

To investigate bargaining under collective risk, we consider exactly these two distributional equity factors because they have also been identified as important in previous studies: the distribution of resources to pay for climate change mitigation (*capacity*) and the distribution of the negative effects of climate change (*vulnerability*). Beyond highlighting the crucial importance of ever-increasing collective risk, the 2018 IPCC report reveals the urgent need to understand the effects of differential vulnerability on bargaining. To do so, models of climate change bargaining must pay more attention to collective risk. The model introduced in this study lends insight into how bargaining behavior might respond to inequalities in capacity and vulnerability during bargaining under collective risk.

What insights can bargaining theory offer for negotiations under increasing collective risk? In bargaining theory, each actor's willingness to accept a proposed agreement depends not only on the terms of the agreement itself, but also on available alternatives. Actors consider the expected value of rejecting a proposal and choosing, instead, to continue bargaining. Central to standard alternating-offer bargaining games is the *continuation value*, or an actor's expected value of continuing to bargain rather than accepting a proposed offer.

The success of an international climate agreement therefore depends on whether parties find it preferable to continue bargaining, and the decision to continue bargaining in this context would necessarily rely heavily on actors' collective risk. In addition to their risk of loss (i.e. vulnerability), the value of continuing to bargain also depends on the how much actors stand to lose (i.e. their capacity). Capacity and vulnerability are thus the key factors in determining continuation values in the context of climate bargaining. As such, these are the two factors that we manipulate in our experimental design, which uses the climate bargaining model with collective risk developed here as a framework. Though wealth is not perfectly correlated with capacity to contribute to climate change mitigation, it is a very good proxy. In the experiment described below the entirety of each player's wealth can be used to mitigate climate. In effect, capacity is exactly proportional to a player's wealth.

Given the highly unequal distribution of capacity and vulnerability across countries and individuals, these two factors are especially important for a better understanding of bargaining behavior under conditions of increasing collective risk. As evidence of their importance, the factors form the basis for determining the organization of regional and geographic groups engaged in negotiations under the United Nations Framework Convention on Climate Change (UNFCCC).[1] The Umbrella group, composed of developed countries, and the Least Developed Countries group are both defined by their level of economic development (capacity) whereas the Small Island Developing States and the Vulnerable 20 (V20) represent overlapping sets of countries with both low capacity and high vulnerability.

These distributional equity factors are not merely academic curiosities. Rather, concerns about capacity and vulnerability feature prominently the official negotiating

[1]The UNFCC is the United Nations body that oversees climate negotiations. Like many negotiations in the United Nations (UN), those in the UNFCCC occur across many levels of regional and geographic groups.

positions of the parties for the 2015 Paris negotiations. In Paris, each country submitted a plan of action to combat climate change in the form of "Intended Nationally Determined Contributions" (INDCs). The importance of capacity and vulnerability in climate negotiations and the difficulty in disentangling the two is evident from Kiribati's INDC (Republic of Kiribati 2016):

> As one of the most vulnerable countries in the world to the effects of climate change [Kiribati's] ability to respond to climate risks is hampered by its highly vulnerable socio-economic and geographical situation.

Despite the challenges posed by global inequalities in collective risk, without an international climate agreement, it will continue to increase for all parties — even if at any point in time some parties are more vulnerable than others. Consequently, each failure to reach a mitigation agreement ensures that subsequent agreements will be negotiated under a higher level of collective risk. There is an extensive literature on collective risk using a public goods framework, particularly the threshold public goods game known as the collective risk social dilemma (Dreber and Nowak 2008; Milinski, Sommerfeld, et al. 2008). While Gampfer (2014), Gosnell and Tavoni (2017), and Smead et al. (2014) use bargaining games to model mitigation efforts, scholars have not yet incorporated the crucial concept of increasing collective risk into bargaining frameworks.

To model mitigation behavior, we develop a "climate bargaining game" that embeds collective risk in an alternating-offers bargaining framework. continuation values play an important role in standard alternating-offers bargaining games[2] so it is imperative to understand how continuation values are affected by the presence of collective risk. Despite the profusion of bargaining models in the international relations literature (Reiter 2003), virtually no bargaining models, barring those mentioned, have been applied climate negotiations over mitigation, leading several prominent political scientists to call for more research on the politics of international climate change (Javeline 2014; Keohane 2015). Here we heed this call by introducing a novel and general model of climate change mitigation bargaining and an experimental framework for investigating its implications. Our framework can be thought of as a generalization of the modified Ultimatum game in Gampfer (2014), in the same way that Rubinstein bargaining (Rubinstein 1982) is a generalization of the Ultimatum game.

[2] The breakthrough in Rubinstein (1982) was the proof of the uniqueness of a sub-game perfect Nash equilibrium (Avery and Zemsky 1994), wherein actors settle on an agreement in the first period due to bargaining costs or a discount factor that reduces their continuation values. Subsequent modifications to Rubinstein's game that explore the robustness of the equilibrium (Shaked et al. 1987; Shaked and Sutton 1984) or how the equilibrium changes under different decision-making procedures (Baron and Ferejohn 1989) all rely on adjustments that alter actors' continuation values.

In a standard alternating offers framework, continuation values depend on actors' individual discount rates and the payoffs foregone by failing to reach an agreement. In the context of bargaining under the shadow of collective risk, continuation values depend not only on the amount that players stand to lose (i.e., their capacity), and their discount rates but also on collective risk, or players' risk of loss from unmitigated climate change (i.e., their vulnerability). As a result, we expect that continuation values should play an even more important role in determining behavior in such a context. It is therefore important to understand whether and how bargaining behavior under collective risk reflects these equity considerations.

In addition to capacity and vulnerability, countries' willingness to engage in burden-sharing also depends on their historical *responsibility* for climate change. For rational egoistic actors (hereafter "rational actors"), however, responsibility only impacts continuation values insofar as it increases the capacity of responsible actors at the onset of climate bargaining.[3] Nonetheless, because responsibility is a mainstay of current climate equity discussions and central to many countries' INDCs, we investigate the effects of varying it, in settings with and without differential vulnerability. We find that responsibility has minimal effects on subjects' behavior above and beyond the direct, mechanical effects it has on continuation values. Additionally, the responsibility treatment also allows us to replicate the control and vulnerability treatments in settings where subjects' starting capacities (and thus, their continuation values) are identical to those in their counterpart treatments except for being the products of endogenous choices. This offers something of a placebo test for the importance of continuation values in this bargaining context. For clarity and brevity, we relegate further discussion of responsibility to Appendix E and devote the remaining discussion to the capacity and vulnerability manipulations.

Given the theoretical importance of continuation values in our bargaining framework, we use them as the strategic primitive and basis for behavioral predictions about how capacity and vulnerability affect bargaining behavior. The results of our experiments substantiate the importance of continuation values, which determine both the size of the offers and the likelihood of success in the bargaining game. The findings suggest the need for further studies of collective risk in the context of climate bargaining.

Despite the United States' recent decision to withdraw from the Paris Climate Accord, the agreement still represents the most ambitious international climate change agreement to date (Davenport 2015). However, for countries to adhere to the commitments they made in Paris, domestic political will must be marshaled and sustained. Moreover, these commitments must be renegotiated indefinitely every five

[3]The assumption of rational egoism entails that actors are self-interested and, in the context of this experiment, motivated by monetary payoffs made to participants over and above a show-up fee (refer to Section 3 for additional details).

years. Thus, understanding what determines states' willingness to bear the costs of mitigation in light of collective risk has never been a more important task for political and social science. Our experimental bargaining framework allows us to effectively manipulate key distributional criteria in climate negotiations. Our results may even suggest some ways of highlighting or downplaying particular distributional equity factors to persuade the United States to re-engage in international climate negotiations.

Following common practice in research experiments, we use a sample of university students in this study. Investigating the behavior of students and other laypersons to better understand the relevant strategic and equity considerations in international climate negotiations is useful and important for several reasons. In the post-Paris world, public preferences and public opinion regarding the willingness to pay for climate change mitigation may be more important than ever before. In the scholarship on public opinion toward climate change, there is evidence for a "bottom up" process through which elite behavior is affected by public preferences (Bechtel and Scheve 2013; Tingley and Tomz 2014). Thus, even if the mass-public is not in a position to directly implement their preferred bargaining strategies, understanding the factors that drive such preferences is important in its own right.In fact, the preferences of the mass public and state and local governments in the United States have already proven to be a powerful counterweight to the Trump Administration's decision to withdraw from the Paris Accord.[4]

The structure of the paper is as follows. Section 2 describes our theoretical framework and how it builds on previous research. Section 3 presents our experimental design and Section 4 presents the results of our four treatments. Section 5 concludes by highlighting the core contributions of the paper and suggests extensions for future work.

2 A climate bargaining game

In the political economy literature, a common way to model climate change mitigation behavior is through the use of public goods games. Many employ variations of standard linear public goods games (e.g. Barrett and Dannenberg 2012, 2014; Hasson, Löfgren, and Visser 2010, 2012, among many others). Other public goods

[4]Even under the extreme assumption that citizens' preferences and public opinion have no influence over international climate negotiations, investigating the behavior of citizens (in our case students) is nonetheless fruitful, for at least two reasons. First, there is some evidence that average behavior in student samples is observationally equivalent to that of elites (Hafner-Burton et al. 2014), and elites exhibit many of the same biases in decision making as the public (Sheffer et al. 2018). Second, as long as a treatment effect is homogeneous across different groups of the population (e.g., students versus elites) any convenience sample will uncover an unbiased estimate of such a (homogeneous) treatment effect (Druckman and Kam 2011).

models have been created specifically for the climate change context. Among them is the collective risk social dilemma (CRSD), a behavioral political economy framework that incorporates the notion of catastrophic economic loss in the event of a failure of mitigation efforts, i.e., collective risk (Dreber and Nowak 2008; Milinski, Sommerfeld, et al. 2008). The CRSD is a threshold public goods game of loss avoidance: players are randomly assigned to small groups and exogenously endowed with an initial wealth level, portions of which can be contributed over multiple rounds toward an exogenously given threshold value. If the group's collective contributions meet or exceed this threshold, the group avoids a collective loss of earnings and all members retain the un-contributed remainder of their endowment. If the group's collective contributions fail to reach the threshold value, the members face an exogenous and known probability of climate change-induced "catastrophic loss" (loss of the entirety of their retained earnings), otherwise known as collective risk. The social dilemma arises, as in all public goods games, from the incentive to free-ride in hopes that one's group will meet the threshold through the contributions of others.

Capacity, in the form of wealth inequalities, has been explored in a CRSD framework using both behavioral game theory and computational simulation (Abou Chakra and Traulsen 2014; Brown and Kroll 2017; Burton-Chellew, May, and West 2013; Milinski, Röhl, and Marotzke 2011; Tavoni et al. 2011; Vasconcelos et al. 2014; J. Wang, Fu, and L. Wang 2010). Kline et al. (2018) and Del Ponte et al. (2017) introduce endogenous (causal) responsibility into the CRSD by making the cost of climate change mitigation and the probability of loss endogenous to group wealth levels that are generated in a preceding common pool resource game. Waichman et al. (2014) investigate asymmetric vulnerability in a CRSD framework, finding that it increases cooperation.

Here we introduce the climate bargaining game (CBG), a modified alternating-offers bargaining framework that incorporates the important feature of increasing collective risk. Players begin the bargaining phase with an initial endowment, which may differ in amount (the *capacity* conditions).[5] They must then bargain over how to split the cost of climate change mitigation. The players bargain under collective risk, therefore failure to reach an agreement exposes both players to catastrophic economic loss. If catastrophic economic loss occurs, each player loses his/her endowment with a predetermined probability that increases monotonically in the number of rejected offers. The players' initial probabilities of loss and the rate at which they increase are, in some cases, asymmetric — the *vulnerability* conditions.

Continuation values represent the payoff that a player could expect to receive should a bargain fail to be reached in the current round. A larger continuation value implies a greater incentive to prolong the bargaining and therefore greater bargaining

[5]In the *responsibility* conditions in Appendix E, the endowment also differs in the manner by which it was obtained.

power. In Rubinstein bargaining, relative "patience" increases one's bargaining power. By discounting the future at a lower rate than their counterparts, patient actors increase their continuation values and, in turn, decrease their willingness to accept low offers. Similarly, in the CBG, because they can afford to be more patient, less vulnerable actors have larger continuation values and therefore greater bargaining power. Higher capacity also translates into larger continuation values and therefore greater bargaining power.

As discussed above, by manipulating endowments (*capacity*) and risk (*vulnerability*) we directly affect continuation values. Table 1 describes the treatments associated with varying each of the two equity dimensions. In the next sections, we use the continuation values determined by our manipulation of endowments and risk schedules to derive predictions about the sizes of offers and the probability of successful bargaining under collective risk. Notably, existing research applying bargaining frameworks to climate change mitigation do not yet consider collective risk (Gampfer 2014; Gosnell and Tavoni 2017; Smead et al. 2014)[6]

3 Experimental design and treatments

For this experiment, a sample of 182 subjects was recruited from the general undergraduate population at a public university in the Northeast United States. The subjects played variations of the CBG, described more fully in the following sections. The experimental sessions were fully computerized. Participants were paid a show-up fee in addition to payments from the experiment itself. A summary of subject demographics is provided in Appendix A.[7]

3.1 General model

The CBG has three parts: the allocation of initial wealth and collective risk schedules; the determination of the cost of climate change mitigation; and the bargaining phase with alternating offers and uncertainty with respect to catastrophic loss. We are able to isolate each of these two key distributional equity factors by manipulating parameters across a number of distinct treatment conditions.

All treatments begin with two randomly and anonymously matched subjects, Players A and B. Then, each player receives an exogenous initial endowment level, which is common knowledge. After the initial endowment is awarded, players use alter-

[6]Gosnell and Tavoni (2017) includes increasing mitigation costs, but not collective risk, in their framework.

[7]Across treatments, the distributions of subjects' gender, age, and political affiliations are similar. Additionally, subjects hold similar opinions on direction and magnitude of the effects of global warming.

nating offers to bargain over how to split mitigation costs. This is meant as a very rough, dyadic approximation of the bargaining process outlined in the Paris Agreement in 2015, in which each county put forth their "offer" as an INDC. The Paris agreement calls for these INDCs to be re-negotiated every five years. The total cost of mitigation to be divided between the players is equal to half the sum of their initial endowments. If an offer is accepted, the match ends, and each player pays the amount agreed upon in the accepted proposal. After each rejection, there is a constant (0.25) probability that the match ends. If this termination is realized, then the players' collective risk is activated, and there is some positive probability that each player in the pair loses the entirety of their endowment. The termination probability and the collective risk are, however, otherwise independent of one another. So, for example, the 0.25 chance that the game ends could be realized, but then whether the players lose their earnings (the collective risk) will be independently determined based on each players' vulnerability. This collective risk, which may differ across players in each match, increases after each rejection. In all treatments, players are randomly re-matched for a total of eight matches. In the description that follows below, the values of the initial endowments, the costs, and the relative collective risk levels are common knowledge to the subjects in each condition.

In all treatments, Player A makes the first offer of how to allocate mitigation costs between the two players. Player B can then either accept or reject the offer. In case of acceptance in a given round t, each player pays their agreed-upon portion of the costs of climate change mitigation and each player's payoff is the difference between their initial endowment and their portion of the accepted allocation of costs.

In the event of rejection, the game continues with probability 0.75. If the first draw from the continuation probability distribution dictates that the game ends, then each player's expected payoff is their initial endowment multiplied by their respective risk after the t'th rejection.

If the game continues, Player B is then able to make a counter-offer and then Player A is given the opportunity to accept or reject it. The procedures for and consequences of acceptance and rejection are the same as in the first round, with the exception of the values of collective risk. Such a loss is possible in the CBG as long as players fail to reach an agreement to split the costs of climate mitigation efforts, which would prevent it. The longer that agreement on the distribution of costs of climate change mitigation is delayed, the greater the collective risk. The probability of loss is exogenously determined as a function of the bargaining history and the treatment assignment. In a particular match, as the tally t of rejected offers increases, so does the risk of loss. For a given player, the probability increases by the same amount after each rejection. In all conditions, we allow a maximum of eight rejected offers, so $t = [1, 2, 3, 4, 5, 6, 7, 8]$. Though the increase in the probability of catastrophic loss varies across conditions and players, it is always positive, and therefore the probability of loss is always strictly monotonically increasing in the number of rejected offers.

We operationalize differential **capacity** by manipulating the initial endowments, such that the initial endowment of Player A is greater than the initial endowment of Player B. Manipulations of each player's collective risk and the amount by which it increases represent differential **vulnerability**. Here, we consider the case in which Player B's probability of catastrophic loss is higher than that of Player A and increases by a greater amount after each rejected offer. Player B, then, is more vulnerable than Player A. Next we describe each of these manipulations in detail.

3.2 Treatments

We are interested in the effects of manipulating initial endowments and collective risk, which result in asymmetric capacities and vulnerabilities between Players A and B.

The baseline treatment (BL) considers bargaining among actors with the same capacities and vulnerabilities, providing a baseline against which we can compare our treatment conditions. Initial endowments are fixed at 100 for each player, resulting in climate change mitigation costs of 100. The BL value for the probability of catastrophic loss for both players is 0.2 for each player and increases by 0.1 after each rejected offer. If all eight possible offers were rejected, then each player's probability of catastrophic loss at the end of the match would be 0.9 and expected payoffs for each player would be amount to 90 percent of the initial endowment. The collective risk schedule[8] over time for each player under BL is (0.2, 0.3, ...0.9). Recall, however, that even if an offer is rejected in round t, the associated collective risk is only activated if the game ends, which occurs with probability 0.25 if the offer is rejected. If the end of the game is reached due to rejection in round t, the expected payoff for each player is the probability of catastrophic loss multiplied by the initial endowment. If the end of the game is reached as a result of an accepted offer, the payoff for a player is the difference between their initial endowment and what they agreed to pay under the proposed offer.

In the baseline treatment, the players are symmetric, therefore Players A and B are undifferentiated. The treatments to follow are all asymmetric, with player A always advantaged either in terms of capacity (wealth) or vulnerability (the likelihood of losing one's wealth), or both. For the sake of consistency, Player A, as the advantaged player, always makes the first offer. This is meant to capture the notion that high income countries tend to be the agenda setters in international negoatiations.

Capacity considerations are incorporated by introducing heterogeneity in initial en-

[8]The collective risk schedule reflects the risk of catastrophic loss in every round.

dowments such that the initial endowment of Player A is greater than the initial endowment of Player B. This gives A greater capacity to contribute to climate change mitigation than B. In this case, Player A's initial endowment is 150 and Player B's initial endowment is 50. Otherwise, the **capacity** (CP) treatment is identical to the BL condition. Comparing results from BL with CP captures the effect of *capacity* considerations in climate change mitigation, allowing us to isolate the effect of "ability to pay" considerations on bargaining under increasing collective risk.

Vulnerability considerations are incorporated by introducing heterogeneity in each player's collective risk. Player B's risk is greater than that of Player A, implying that B is more vulnerable to the effects of climate change (catastrophic economic loss) than A, and B's vulnerability increases more quickly than A's. As in the BL condition the initial endowment for each player in the basic **vulnerability** (VN) treatment is 100. The respective initial risks for Players A and B in the VN treatment, however, are 0.05 and 0.2, and Player B's probability of catastrophic loss increases more rapidly than that of Player A. After each rejection, Player A's risk increases by 0.05, whereas Player B's risk increases by 0.1. In VN, the collective risk schedules for Player A and Player B respectively are (0.05, 0.10, ...0.4) and (0.2, 0.3, ...0.9). By comparing the VN conditions to BL, we can determine what weight the subjects place on the distribution of *vulnerability* to the impacts of climate change when making their offers about how mitigation costs should be split. This manipulation allows us to isolate the effect of the "beneficiary pays" principle on bargaining. Note that, relative to BL both the initial level of collective risk and the rate of its increase are being manipulated. While this prevents us from disentangling the effects of each of these changes, the design still allows us to determine the effect of asymmetric vulnerability as it is operationalized in the design.

The capacity-vulnerability (CP-VN) condition, designed to combine the endowment parameters of CP with the risk parameters of VN, investigates the interactive effects of the CP treatment and the VN treatment. As in the CP treatment, Player A's initial endowment is 150 and Player B's initial endowment is 50. Additionally, as in the VN treatment, the collective risk schedules for Player A and Player B respectively are (0.05, 0.10, ...0.4) and (0.2, 0.3, ...0.9). Combining the capacity and vulnerability conditions in this way allows us to observe interactions between capacity and vulnerability concerns when compared to the BL conditions as well as to the simple CP and VN conditions.

A summary of the treatments is provided in Table 1. Further details regarding the dates of administration and instructions given to the subjects are given in Appendix B, and discussion of the responsibility treatments are in Appendix E.

[9]For initial endowments and collective risk, parameters are presented in order for Players A and B.

Table 1: Summary of Treatments

Treatment	**Initial Endowment**[9]	**Collective Risk**	**Number of Subjects**
BL	100/100	(0.2, 0.3, ...0.9) (0.2, 0.3, ...0.9)	42
CP	150/50	(0.2, 0.3, ...0.9) (0.2, 0.3, ...0.9)	44
VN	100/100	(0.05, 0.1, ...0.4) (0.2, 0.3, ...0.9)	48
CP-VN	150/50	(0.05, 0.1, ...0.4) (0.2, 0.3, ...0.9)	48

3.3 Equilibrium analysis

Both wealth and risk of loss factor into the calculation of continuation values, and so continuation values under the *VN*, *CP*, and *CP-VN* differ from those under *BL*. Recall that continuation values describe a strategically rational player's expected payoffs from continuing to bargain rather than accepting a rational counterpart's proposed agreement. In a given round, rational risk-neutral actors accept an offer if and only if it is no less than the player's continuation value.

To calculate continuation values, we use the notation $i \in \{A, B\}$ to denote each player; $t \in \{1, 2, \ldots, 8\}$ for rounds; p_i^t for i's probability of loss in round t; E_i for i's endowment; and V_i^t for i's continuation value in round t. In a given round, V_i^t is the weighted sum of (i) the payoff from the case in which the game ends prematurely and each player facing its probability of loss for that round $0.25 \cdot (1 - p_i^t) \cdot E_i^t$; and (ii) the expected payoff from the game continuing with equilibrium behavior in all subsequent rounds, which is the continuation value for player i in the subsequent round weighted by the probability that the game continues, $0.75 \cdot V_i^{t+1}$. If an offer is rejected *and* the game continues, the responder in t will be the offerer in $t + 1$ and vice versa, the offerer in t will be the responder in $t + 1$. In all conditions, Player A (B) is the offerer (responder) in the first and all subsequent odd rounds, and B (A) is the offerer (responder) in all even rounds. Recall that in the final round $t = 8$, the probability of the game continuing after a rejection is zero and therefore the probability of the game ending is 1.

We calculate player $i = A$'s continuation value in period t as:

$$V_A^t = \begin{cases} \overbrace{.25 \cdot (1 - p_A^t) \cdot E_A^t}^{\text{Game ends}} + \overbrace{.75 \cdot V_{A=R}^{t+1}}^{\text{Game continues with } A \text{ as responder}} & \text{if } t = 1, 3, 5, 7 \\ \overbrace{.25 \cdot (1 - p_A^t) \cdot E_A^t}^{\text{Game ends}} + \overbrace{.75 \cdot V_{A=O}^{t+1}}^{\text{Game continues with } A \text{ as offerer}} & \text{if } t = 2, 4, 6 \\ 1 \cdot (1 - p_A^t) \cdot E_A^t + 0 \cdot V_{A=O}^{t+1} & \text{if } t = 8 \end{cases} \tag{1}$$

Likewise, we calculate player $i = B$'s continuation value in period t as:

$$V_B^t = \begin{cases} \overbrace{.25 \cdot (1 - p_B^t) \cdot E_B^t}^{\text{Game ends}} + \overbrace{.75 \cdot V_{B=O}^{t+1}}^{\text{Game continues with } B \text{ as offerer}} & \text{if } t = 1, 3, 5, 7 \\ \overbrace{.25 \cdot (1 - p_B^t) \cdot E_B^t}^{\text{Game ends}} + \overbrace{.75 \cdot V_{B=R}^{t+1}}^{\text{Game continues with } B \text{ as responder}} & \text{if } t = 2, 4, 6 \\ 1 \cdot (1 - p_B^t) \cdot E_B^t + 0 \cdot V_{B=R}^{t+1} & \text{if } t = 8 \end{cases} \tag{2}$$

Because this is a finitely repeated bargaining game with a known maximum number of rounds, we can use backward induction to calculate continuation value for each of the players, and begin our analysis in the final round with the decision that confronts the responder, player A. The final round is strategically similar to an ultimatum game: the responder in the final round A gets a take-it-or-leave-it offer, with the payoff for the leave-it option being $(1 - p_A^8) \cdot E_A = V_A^8$, the continuation value for player A if they reject the offer in the eighth round. Following equation 1 we can recursively calculate the remaining values by plugging it into the equation $V_A^7 = 0.25 \cdot (1 - p_A^7) \cdot E_A^7 + 0.75 \cdot V_A^8$. Similarly, the continuation value for B if their offer is rejected in the eighth round is $V_B^8 = (1 - p_B^8) \cdot E_B$, and their values can then be recursively calculated following (2).[10]

V_A^8 is the continuation value for the responder in round eight. Player B, the offerer in round eight, therefore knows that the maximum cost the responder would be willing to pay is $C_A^* = E_A - V_A^8$. This is the minimum offer that B can make in the final round that would be accepted by A. In this case the B's implied cost is $C_B^* = 100 - C_A^*$. This alone, however, is not sufficient to determine whether B would be willing to offer C_B^*, and therefore whether there will be an offer accepted in equilibrium. This will only be the case if $C_B^* \leq V_B^8$.

Taking the BL condition as an example, $V_A^8 = (1 - 0.9) \cdot 100 = 10$, thus $C_A^* = 100 - 10 = 90$ and $C_B^* = 100 - 90 = 10$. V_B^8 also equals 10, so $C_B^* \leq V_B^8$, and B's equilibrium offer would be accepted by A.

[10]We would like to thank an anonymous reviewer for guiding us to the solution we are now using.

Table 2 summarizes the parameters relevant for the continuation values in each condition. For ease of presentation, players are separated on the basis of roles (i.e., the probability of loss for the offerer in round t is denoted p^t_{Offerer} and that of the responder is $p^t_{\text{Responder}}$). Note that, in equilibrium, offers are not accepted until the fourth round at the earliest (BL and CP) or the eighth round at the latest (CP-VN). So, if we expect equilibrium play we should not expect to see offers accepted in the first few rounds.

Table 2: Summary of Continuation Values, Equilibrium Offers, and Equilibrium Responses

Treatment	Round	p^t_{Offerer}	$p^t_{\text{Responder}}$	E_{Offerer}	$E_{\text{Responder}}$	V^t_{Offerer}	$V^t_{\text{Responder}}$	C^*_{Offerer}	Accept/Reject
	1	0.2	0.2			58.3	67.1	41.7	Reject
	2	0.3	0.3			62.7	51.0	37.3	Reject
	3	0.4	0.4			44.7	60.3	55.3	Reject
	4	0.5	0.5			60.4	39.6	39.6	Accept
BL	5	0.6	0.6	100	100	36.1	63.9	63.9	Accept
	6	0.7	0.7			71.9	28.1	28.1	Accept
	7	0.8	0.8			27.5	72.5	72.5	Accept
	8	0.9	0.9			90.0	10.0	10.0	Accept
	1	0.2	0.2	150	50	89.9	35.4	60.1	Reject
	2	0.3	0.3	50	150	33.8	79.9	16.2	Reject
	3	0.4	0.4	150	50	71.5	33.5	78.5	Reject
CP	4	0.5	0.5	50	150	34.6	65.4	15.4	Accept
	5	0.6	0.6	150	50	62.2	37.8	87.8	Accept
	6	0.7	0.7	50	150	43.8	56.3	6.3	Accept
	7	0.8	0.8	150	50	60.0	40.0	90.0	Accept
	8	0.9	0.9	50	150	50.0	50.0	0	Accept
	1	0.05	0.2			82.7	58.0	17.3	Reject
	2	0.3	0.1			50.7	78.6	49.3	Reject
	3	0.15	0.4			74.7	44.2	25.3	Reject
	4	0.5	0.2			39.0	71.3	61.0	Reject
VN	5	0.25	0.6	100	100	68.4	35.3	31.6	Reject
	6	0.7	0.3			33.8	66.3	66.3	Accept
	7	0.35	0.8			65.0	35.0	35.0	Accept
	8	0.9	0.4			40.0	60.0	60.0	Accept
	1	0.05	0.2	150	50	123.0	27.7	27.0	Reject
	2	0.3	0.1	50	150	23.6	116.5	26.4	Reject
	3	0.15	0.4	150	50	110.3	19.7	39.7	Reject
$CP-VN$	4	0.5	0.2	50	150	16.3	104.6	33.7	Reject
	5	0.25	0.6	150	50	99.5	13.4	50.5	Reject
	6	0.7	0.3	50	150	11.3	95.2	38.8	Reject
	7	0.35	0.8	150	50	91.9	10.0	58.1	Reject
	8	0.9	0.4	50	150	10.0	90.0	45.0	Accept

4 Results

The experimental manipulation of subjects' wealth (capacity) and collective risk (vulnerability) is reflected in their continuation values, which are central to our analysis. The finite nature of our game and the backward induction invoked to solve it imply that continuation values also depend on the order in which offers are made. Consequently, we focus on differences across conditions, and, where appropriate, across rounds and within player types. In other words, we will often be comparing the behavior of a player type — A or B — across conditions and over time.

We begin with an analysis of the rates of successful bargaining across conditions and rounds. We are interested in how our capacity and vulnerability manipulations affect bargaining behavior under the threat of increasing collective risk. Most fundamentally, we are interested in the effect of our manipulations on successful negotiation. For each condition, Figure 1 shows the cumulative percentage of accepted offers across rounds one through eight. For example, in *BL*, by round three, an acceptable offer had been made in approximately 50 percent of matches, and by round five, this figure had increased to 60 percent.

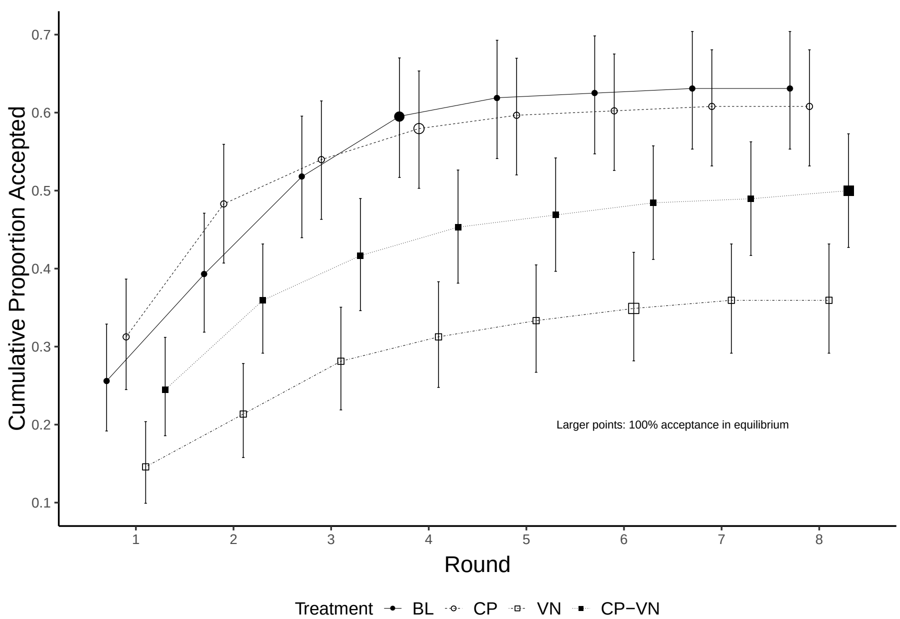


Figure 1: Cumulative Offer Acceptance Rate by Condition

The pattern in Figure 1 raises two puzzles. First, why are the observed rates of success, especially in the early rounds, so much higher than those predicted? Second, why do rates of success differ across conditions, notably between the asymmetric vulnerability conditions (*VN* and *CP-VN*) and their symmetric counterparts (*BL* and *CP*)?

Success across the eight rounds is notably lower in the two asymmetric vulnerability conditions than in their counterpart conditions. Asymmetric capacity does not appear to hamper success, and in the case of *CP-VN* it appears the pairing with the wealth differences to be marginally increasing success relative to *VN*. Also recall from Table 2, however, that in equilibrium, we should not expect *any* agreement until rounds four (in the case of *BL* and *CP*), six (*VN*) or eight (*CP-VN*). Yet, across all conditions, high rates of observed success occur much earlier than theory would suggest. In contrast, the lower success rates in *VN* and *CP-VN* are consistent both with the later arrival of equilibrium acceptance in those two conditions and with Gampfer (2014), which finds, in single shot ultimatum-style games, that offers increase in wealth and vulnerability, as offerers exploit their counterparts' greater comparative vulnerability by offering them less. Our study extends and generalizes the bargaining framework in Gampfer (ibid.), allowing us to assess the effect of increasing collective risk in a multi-round framework.

Why do we observe such high success rates, and how can we explain the variability in success across conditions? Compared to the theory's predictions, are the offerers too generous? Are the responders too willing to accept low offers? For that we need to look beyond the aggregate success rates displayed in Figure 1. In our theoretical framework a "generous" offer, from the offerer's perspective, is simply one that exceeds the equilibrium offer. Predicted rates of success are derived under the assumption that equilibrium offers are chosen. By definition, generous offers violate this assumption, and so they may explain the relatively high rates of success in bargaining shown in Figure 1. The results in Figure 1 indicate less success in the asymmetric vulnerability conditions than in their symmetric counterparts. However, comparing the difference between each player's first average and equilibrium offer across conditions and player types, a different picture emerges.

These differences are shown in Figure 2. In both *VN* and *CP-VN*, Player A makes average offers that significantly exceed the equilibrium offer. Rather than exploiting their partner's vulnerability, the advantaged parties are generous vis-à-vis our theoretical predictions. In contrast, *BL* and *CP* offers do not, on average, differ from those in equilibrium. For Player B, the results are nearly reversed: on average, *B*'s offers in *VN* and *CP-VN* fall short of those in equilibrium, while those in *BL* and *CP* are close to equilibrium predictions.

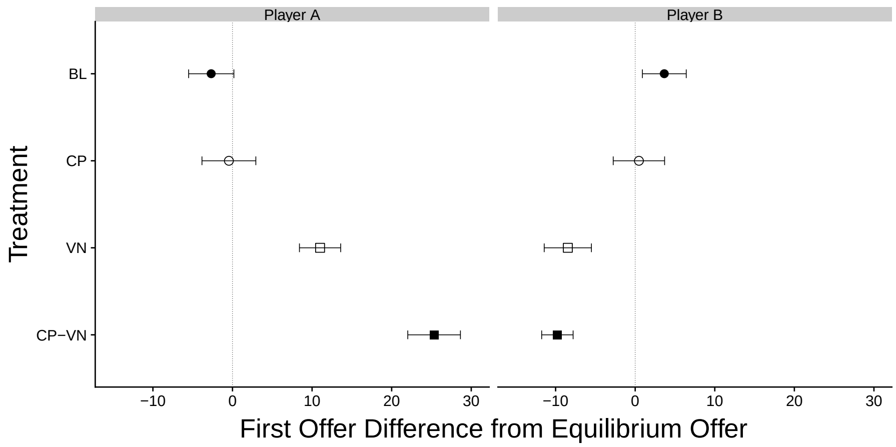


Figure 2: Difference between Observed and Equilibrium First Offers by Type and Condition

Figure 3 displays the trends in the differences presented in Figure 2 over the course of the eight potential rounds. Some interesting patterns emerge. Irrespective of condition, Player A's average offer declines across rounds. While the slopes of trends in players' average offers are approximately equal across the conditions, their intercepts are not. Relative to equilibrium offers, offers in *VN* and *CP-VN* are on average higher in each round than those in *BL* and *CP*. Again, the pattern is approximately reversed for Player B, whose average offers in each round are higher in *BL* and *CP* than in *VN* and *CP-VN*. For *B*, the trends are not parallel. Instead, they diverge as the match progresses. Final round offers in *VN* and *CP-VN* become even less generous and those in *BL* and *CP* become more generous. The probability of acceptance likely increases in the size of offers, and due to this potential for selection across rounds, these effects, though thought-provoking, are not causally identified.

The results displayed in Figure 3 seem to hint at an explanation for the two initial puzzles. Perhaps the relative generosity initially exhibited by Player A in *VN* and *CP-VN* is offset by Player B's relative stinginess in the second round. This pattern of offsetting offers occurs across all eight rounds, revealing interesting treatment effects across types and conditions. It does not, however, seem to explain the differential success rates observed in Figure 1 because the *VN* and *CP-VN* conditions are nonetheless characterized by *lower* rates of success than *BL* and *CP*. So, overall rates of success do not seem to be driven by the offerer's behavior. Perhaps it is the responder, then, whose behavior is driving the puzzling trends?

We have been using the equilibrium offer as our relevant threshold for generosity with respect to the offerers' behavior. To answer the question of whether the responder was too readily willing to accept low offers, we must identify a different metric

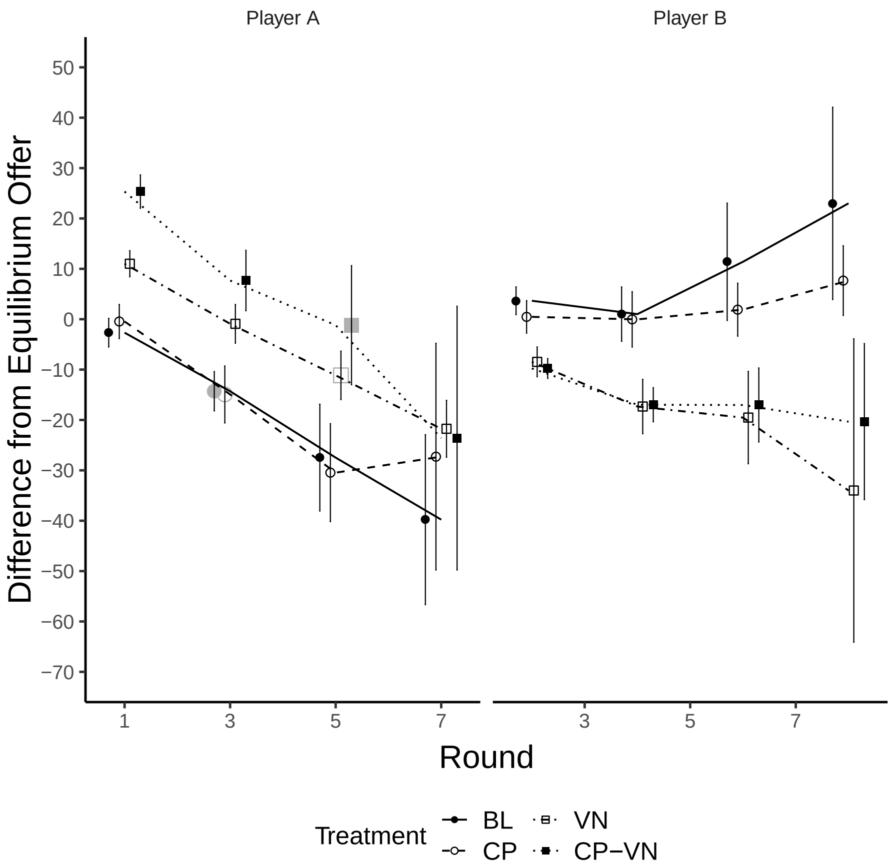


Figure 3: Round-by-Round Differences between Observed and Equilibrium Offers by Type and Condition

that is relevant for the responder's decision to accept or reject. In summarizing the expected payoffs from accepting a bargain, the continuation value serves as a natural reference point for the responder in the alternating offers bargaining game. The theory predicts that they should accept the offer if and only if its net payoff is larger than the expected payoff from rejecting the offer and continuing to bargain. In other words, the responder should only accept the offer if the immediate payoff from accepting the proposed offer, abbreviated as the Net Offer, exceeds the continuation value from continuing to bargain, i.e., if $E_{\text{Responder}} - C_{\text{Responder}} > V_{\text{Responder}}$.

Figure 4 illustrates the distribution of differences between the net offer and the continuation values, separated by round, player, treatment, and the responder's decision to accept or reject the offer. Negative observations below the dotted line describe responses in cases where the continuation value exceeds the net offer, i.e., those in which a rational responder should not accept the offer. It is immediately apparent that the preponderance of offers, including those that were accepted (indicated by the black dots), are below the responder's continuation value, with the highest number of generous offers to Player B made in *BL* and *CP*. Even in those two cases, however, the first two net offers made by B are still largely below A's continuation values. Suprisingly, many are nonetheless accepted by Player A. On the other hand, in *VN* and *CP-VN*, practically none of Player B's offers exceed A's continuation value.

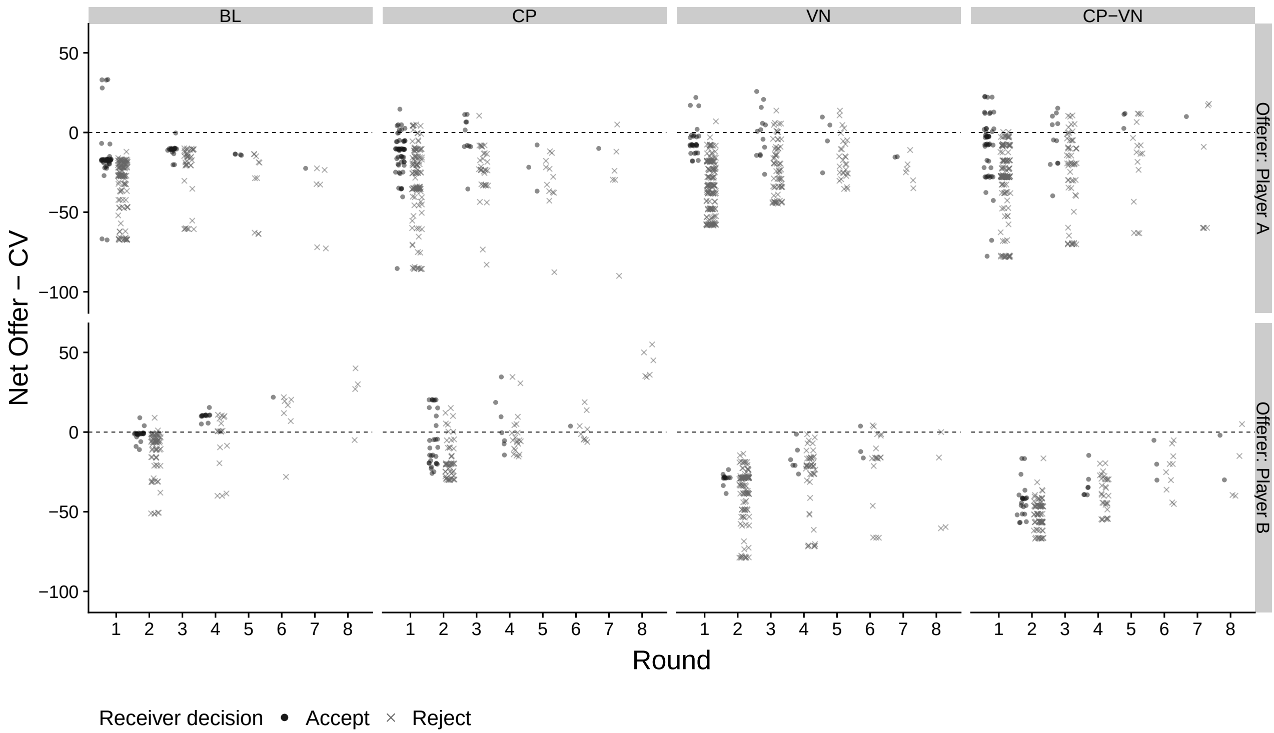


Figure 4: Distribution of Accepted and Rejected Offers by $E_{\text{Responder}} - C_{\text{Responder}} - V_{\text{Responder}}$, Round, Player, and Treatment. The four vertical panels correspond to the four treatments, and rounds are distinguished on the horizontal axis, with odd rounds corresponding to cases where Player A is the offerer (on the top panel) and even ones to those where Player B is the offerer (on the bottom panel). The vertical axis describes the difference between net offers ($E_{\text{Responder}} - C_{\text{Responder}}$) and responders' continuation values ($V_{\text{Responder}}$), with the dashed horizontal line indicating cases where the net offer matches the responders' continuation value. Offers accepted by the responder are illustrated by the black points on the left side of each round, and rejected ones by the grey "x"s on the right. Theoretically, we expect that the receiver will accept net offers greater than the continuation value (above the dotted line) at a higher rate than those below the continuation value.

Across all conditions, Figure 4 includes a large number of accepted offers below the responder's continuation value. In contrast, the behavior of each type of player generally cancels out the other within each condition, as shown in Figure 3. Taken together, these results suggest that responders' willingness to accept relatively low offers — rather than generosity of the offerer — is the major driver of success rates that exceed predictions.

5 Discussion

As the most recent IPCC report makes clear, rapid progress toward mitigation is becoming increasingly urgent if we are to avoid the worst effects of climate climate change (IPCC 2018). In other words, the global collective risk of economic and social catastrophe will continue to increase unless the global community can agree to a plan to prevent it. Such urgency makes understanding behavioral responses to collective risk — a peculiar feature of climate change negotiations — of even greater importance. The IPCC report makes more salient an inescapable truth: the repeated failures of the international community to mitigate climate change are taking place against a backdrop of ever increasing collective risk of inaction. Under the status quo, unless effective mitigation policies are successfully negotiated, collective risk will continue to rise.

Yet collective risk — this very vulnerability to collective inaction — is not evenly distributed across nations or individuals. Nor is the capacity to successfully mitigate such risk. In this study we have developed and tested an alternating offers bargaining game which incorporates increasing collective risk while manipulating the players' vulnerability to collective risk and capacity to mitigate. Analytically, we focus on the concept of a continuation value, or the discounted value of continuing to bargain rather than accepting the status quo offer. Because collective risk directly affects continuation values, they are a proxy for bargaining power. A player with a larger continuation value can more credibly reject offers; they can afford to be more 'patient'. Patience, as manifested by large continuation values, has long been understood as an important determinant of bargaining power (Korobkin 2003; Ponsatí and Sákovics 1998). Populations vulnerable to climate change cannot afford to be as patient in coming to an agreement to solve the problem as those that are less vulnerable. The presence of collective risk in climate bargaining makes patience all the more important in this context. In other words, in international climate negotiations, perhaps even more so than in other situations, patience implies bargaining power.

Intuitively, then, one might expect less vulnerable parties to exploit the more vulnerable parties with low offers, as is the case in the one shot games reported in Gampfer (2014). The results from our alternating offers game, however, do not support such a conclusion. In fact, the less vulnerable parties are more generous compared to a baseline symmetric vulnerability condition. On the other hand, the more vulnerable parties are less likely to accept generous offers by the less vulnerable parties, and respond with even lower offers of their own. In our study at least, it would behoove the more vulnerable parties to be more willing to compromise. Perhaps their reluctance to do so is because they do not want their vulnerability to be exploited. Future research should explore additional mechanisms that might affect these equity and procedural justice considerations. Armed with this knowledge, we can derive guidance about how best to frame climate negotiations to harness these strategic motivations and increase the likelihood of successful negotiations.

As we discussed in the introduction we believe that many criticisms of student samples are misplaced. Still, future research on individual preferences for climate equity should focus on how public attitudes might prevent or enable countries to meet their commitments under the Paris agreement. Such studies could do so by combining representative survey experiments with behavioral games, as suggested in Mutz (2011, Chapter 5). Further research into elite behavior in this context, though difficult, would be welcome, and could give us an empirical answer to the question of whether in this domain the attitudes and behaviors of citizens and elites are in fact similar.

As made clear in IPCC (2018), it has never been more urgent to understand climate change bargaining under increasing collective risk. In fact, collective risk itself makes this even more urgent. We believe that the general framework we introduce here — that of bargaining under collective risk — is as important as any specific finding in this study because it offers a promising tool for us to understand how collective risk affects international climate negotiations. Our study is the first to incorporate this aspect of international climate negotiations into a behavioral model, but we believe that much work remains to be done to fully exploit the flexibility of this general framework.

References

Abou Chakra, Maria and Arne Traulsen (2014). "Under High Stakes and Uncertainty the Rich Should Lend the Poor a Helping Hand". In: *Journal of Theoretical Biology* 341, pp. 123–130.

Alley, Richard B et al. (2003). "Abrupt Climate Change". In: *Science* 299.5615, pp. 2005–2010.

Avery, Christopher and Peter B Zemsky (1994). "Money Burning and Multiple Equilibria in Bargaining". In: *Games and Economic Behavior* 7.2, pp. 154–168.

Baron, David P and John A Ferejohn (1989). "Bargaining in Legislatures." In: *American Political Science Review* 83.04, pp. 1181–1206.

Barrett, Scott and Astrid Dannenberg (2012). "Climate Negotiations Under Scientific Uncertainty". In: *Proceedings of the National Academy of Sciences* 109.43, pp. 17372–17376.

– (2014). "Sensitivity of Collective Action to Uncertainty About Climate Tipping Points". In: *Nature Climate Change* 4.1, pp. 36–39.

Bechtel, Michael M and Kenneth F Scheve (2013). "Mass Support for Global Climate Agreements Depends on Institutional Design". In: *Proceedings of the National Academy of Sciences* 110.34, pp. 13763–13768.

Brown, Thomas C and Stephan Kroll (2017). "Avoiding an Uncertain Catastrophe: Climate Change Mitigation Under Risk and Wealth Heterogeneity". In: *Climatic Change* 141.2, pp. 155–166.

Burton-Chellew, Maxwell N, Robert M May, and Stuart A West (2013). "Combined Inequality in Wealth and Risk Leads to Disaster in the Climate Change Game". In: *Climatic Change* 120.4, pp. 815–830.

Davenport, Coral (2015). "Nations Approve Landmark Climate Accord in Paris". Accessed on August 13, 2017. URL: https://www.nytimes.com/2015/12/13/world/europe/climate-change-accord-paris.html.

Del Ponte, Alessandro et al. (2017). "Passing It Along: Experiments on Creating the Negative Externalities of Climate Change". In: *The Journal of Politics* 79.4, pp. 000–000.

Dietz, Simon et al. (2016). "Climate Value at Risk of Global Financial Assets". In: *Nature Climate Change*.

Dreber, Anna and Martin A Nowak (2008). "Gambling for Global Goods". In: *Proceedings of the National Academy of Sciences* 105.7, pp. 2261–2262.

Druckman, James N and Cindy D Kam (2011). "Students as Experimental Participants". In: *Cambridge Handbook of Experimental Political Science*. Ed. by James N. Druckman et al. Cambridge University Press. Chap. 4, pp. 41–57.

Gampfer, Robert (2014). "Do Individuals Care About Fairness in Burden Sharing for Climate Change Mitigation? Evidence From a Lab Experiment". In: *Climatic Change*, pp. 1–13.

Gosnell, Greer and Alessandro Tavoni (2017). “A Bargaining Experiment on Heterogeneity and Side Deals in Climate Negotiations”. In: *Climatic Change* 142.3-4, pp. 575–586.

Hafner-Burton, Emilie M et al. (2014). “Decision Maker Preferences for International Legal Cooperation”. In: *International Organization* 68.04, pp. 845–876.

Hasson, Reviva, Åsa Löfgren, and Martine Visser (2010). “Climate Change in a Public Goods Game: Investment Decision in Mitigation Versus Adaptation”. In: *Ecological Economics* 70.2, pp. 331–338.

– (2012). “Treatment Effects of Climate Change Risk on Mitigation and Adaptation Behaviour in an Experimental Setting”. In: *South African Journal of Economics* 80.3, pp. 415–430.

IPCC (2018). “Summary for Policymakers”. In: *An IPCC Special Report on the Impacts of Global Warming of 1.5 Degress C Above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*. Ed. by V. Masson-Delmotte et al. Geneva, Switzerland: World Meteorological Organization.

Javeline, Debra (2014). “The Most Important Topic Political Scientists Are Not Studying: Adapting to Climate Change”. In: *Perspectives on Politics* 12.02, pp. 420–434.

Keohane, Robert O (2015). “The Global Politics of Climate Change: Challenge for Political Science”. In: *PS: Political Science & Politics* 48.01, pp. 19–26.

Kline, Reuben et al. (2018). “Differentiated responsibilities and prosocial behaviour in climate change mitigation”. In: *Nature Human Behaviour* 2.9, p. 653.

Korobkin, Russell (2003). “Bargaining Power as Threat of Impasse”. In: *Marq. L. Rev.* 87, p. 867.

Milinski, Manfred, Torsten Röhl, and Jochem Marotzke (2011). “Cooperative Interaction of Rich and Poor Can Be Catalyzed by Intermediate Climate Targets”. In: *Climatic Change* 109.3-4, pp. 807–814.

Milinski, Manfred, Ralf D Sommerfeld, et al. (2008). “The Collective-Risk Social Dilemma and the Prevention of Simulated Dangerous Climate Change”. In: *Proceedings of the National Academy of Sciences* 105.7, pp. 2291–2294.

Mutz, Diana C (2011). *Population-Based Survey Experiments*. Princeton University Press.

Ponsatí, Clara and József Sákovics (1998). “Rubinstein Bargaining With Two-Sided Outside Options”. In: *Economic Theory* 11.3, pp. 667–672.

Reiter, Dan (2003). “Exploring the Bargaining Model of War”. In: *Perspectives on Politics* 1.01, pp. 27–43.

Republic of Kiribati (2016). “Intended Nationally Determined Contribution”. URL: https://www4.unfccc.int/sites/ndcstaging/PublishedDocuments/Kiribati%20First/INDC_KIRIBATI.pdf.

Rubinstein, Ariel (1982). “Perfect Equilibrium in a Bargaining Model”. In: *Econometrica: Journal of the Econometric Society*, pp. 97–109.

Schellnhuber, Hans Joachim (2006). *Avoiding Dangerous Climate Change*. Cambridge University Press.

Shaked, Avner et al. (1987). *Opting Out: Bazaars Versus” Hi Tech” Markets*. Tech. rep. Suntory, Toyota International Centres for Economics, and Related Disciplines, LSE.

Shaked, Avner and John Sutton (1984). “Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model”. In: *Econometrica: Journal of the Econometric Society*, pp. 1351–1364.

Sheffer, Lior et al. (2018). “Nonrepresentative representatives: an experimental study of the decision making of elected politicians”. In: *American Political Science Review* 112.2, pp. 302–321.

Smead, Rory et al. (2014). “A Bargaining Game Analysis of International Climate Negotiations”. In: *Nature Climate Change* 4.6, pp. 442–445.

Tavoni, Alessandro et al. (2011). “Inequality, Communication and the Avoidance of Disastrous Climate Change in a Public Goods Game”. In: *Proceedings of the National Academy of Sciences USA* 108, pp. 11825–11829.

Tingley, Dustin and Michael Tomz (2014). “Conditional Cooperation and Climate Change”. In: *Comparative Political Studies* 47, pp. 344–368.

Vasconcelos, Vítor V et al. (2014). “Climate Policies Under Wealth Inequality”. In: *Proceedings of the National Academy of Sciences*.

Waichman, Israel et al. (2014). “Asymmetry Enhances Success Chances for International Climate Change Negotiations”. In: *WCERE in Istanbul*.

Wang, Jing, Feng Fu, and Long Wang (2010). “Effects of Heterogeneous Wealth Distribution on Public Cooperation With Collective Risk”. In: *Physical Review E* 82.1, p. 016102.

Weitzman, Martin L (2011). “Fat-Tailed Uncertainty in the Economics of Catastrophic Climate Change”. In: *Review of Environmental Economics and Policy* 5.2, pp. 275–292.

Appendix A Distribution of Subject Demographics and Opinions

As shown in Figures 5 through 9, the distribution of subject demographics across treatments was similar. Overall, more women than men participated in the experiments, particularly for the vulnerability and capacity treatments. The distribution across ages was also similar. As expected in the case of average undergraduates, most subjects across the treatments were concentrated between ages 20 and 22. Per Figures 7 through 9, subjects were generally twice as likely to affiliate with the Democratic party as the Republican party (in the case of the uneven probability endowment, nearly three-times as likely) and found global warming to be a realistic threat.

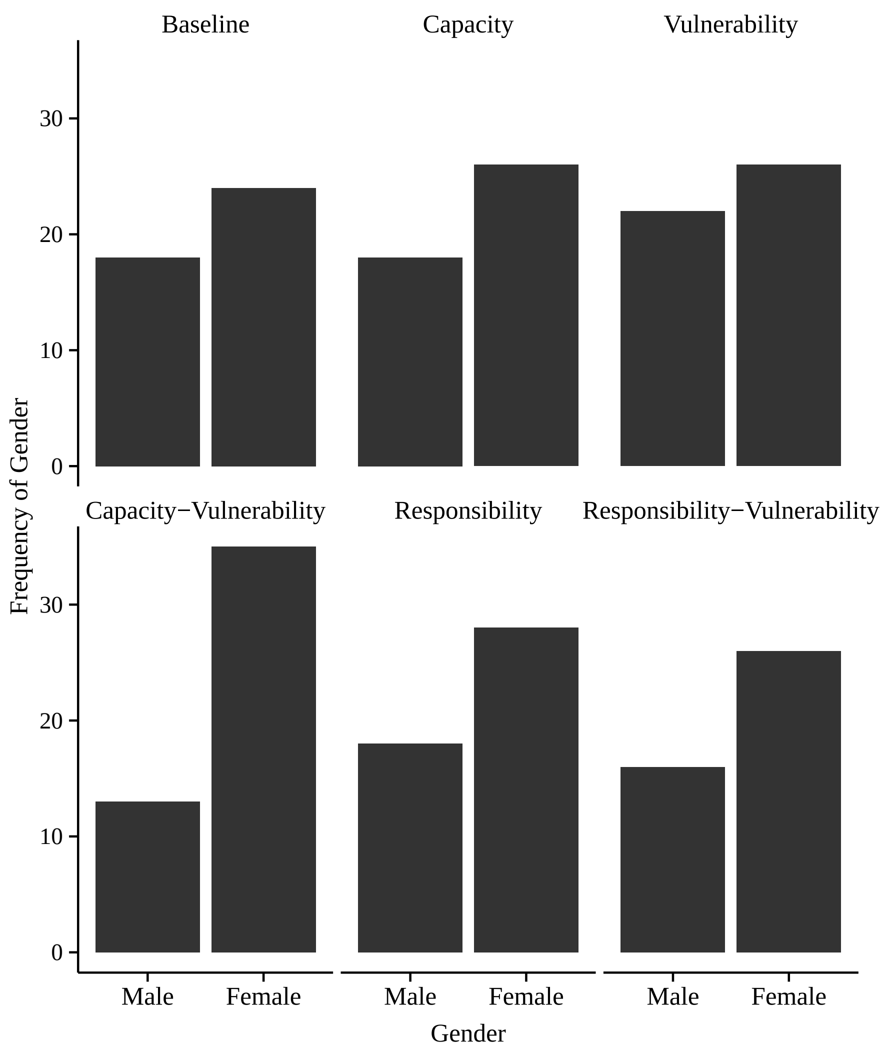


Figure 5: Distribution of Subject Gender

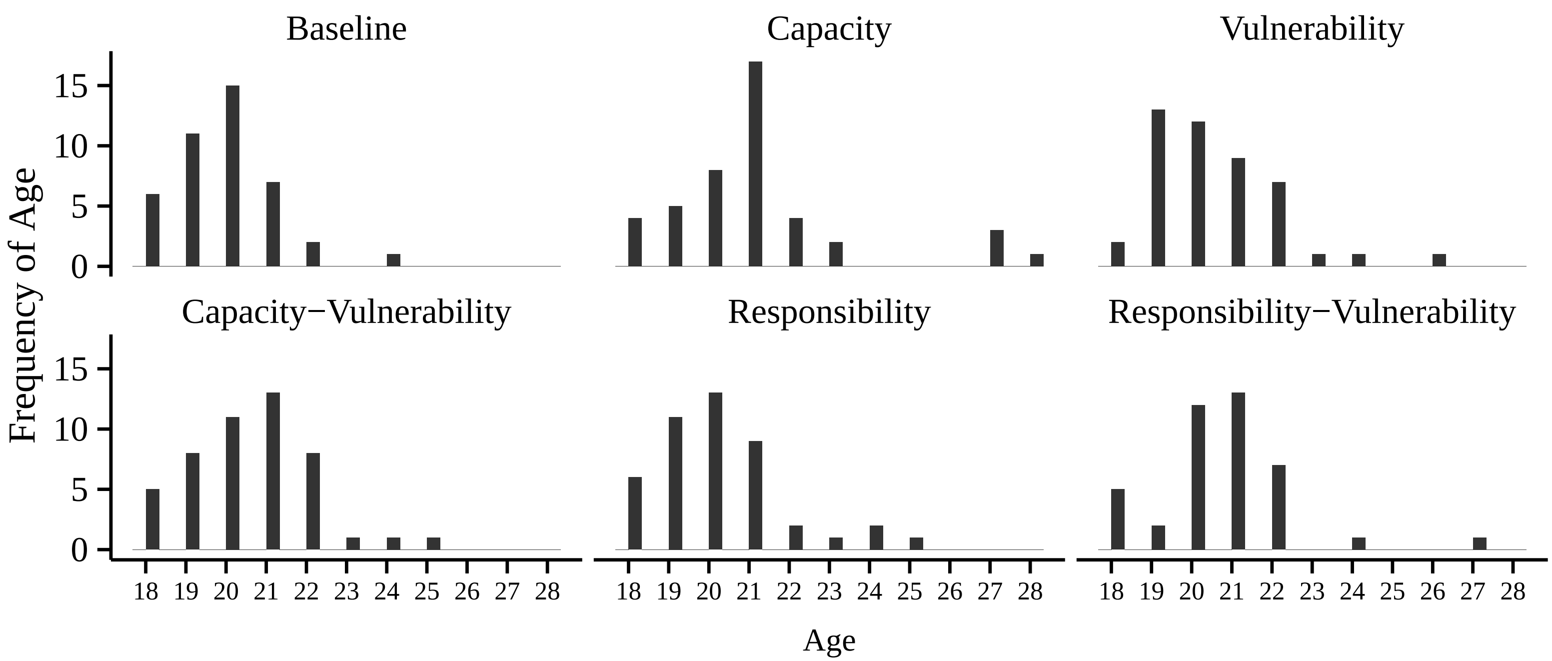


Figure 6: Distribution of Subject Ages

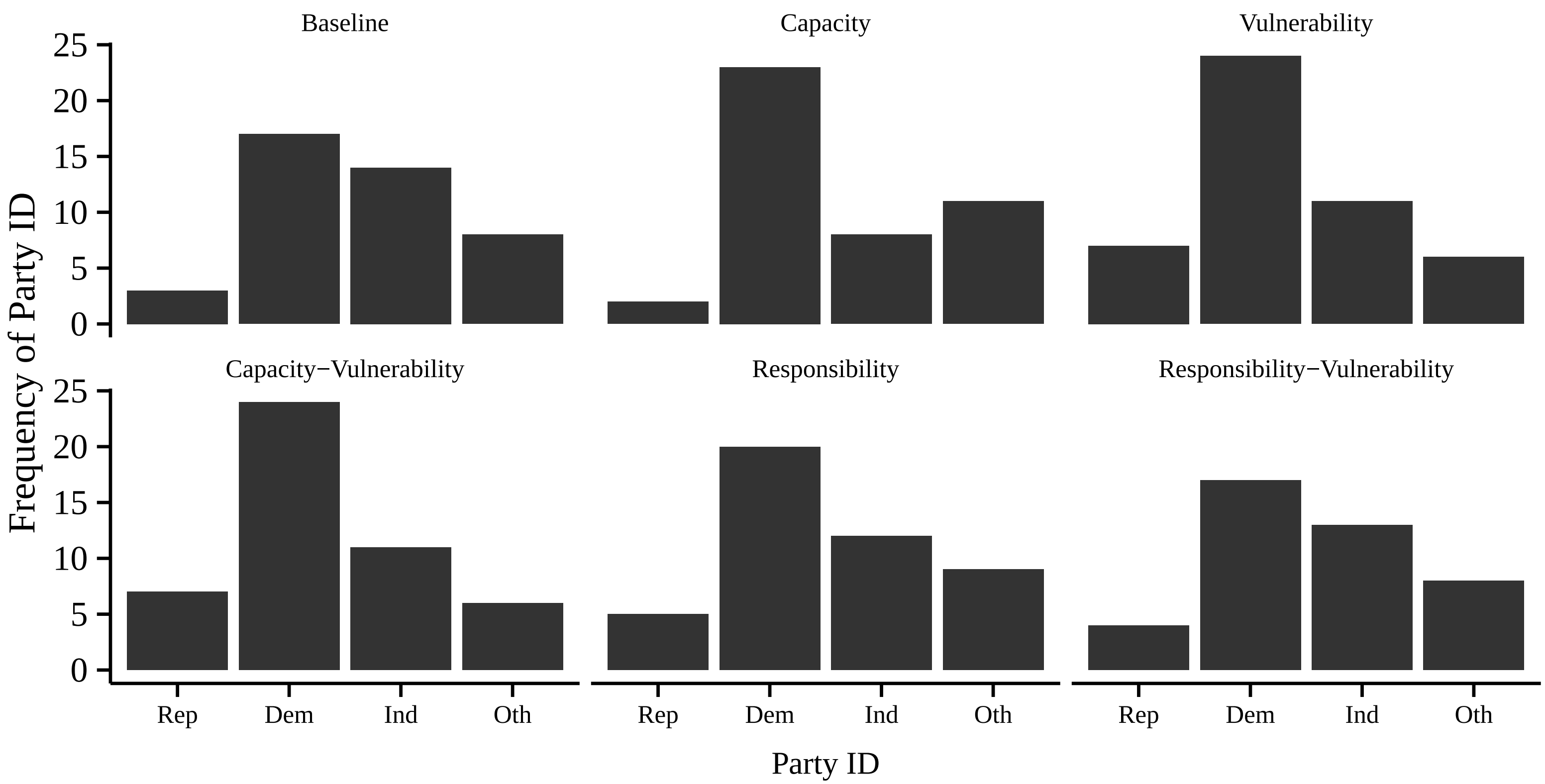


Figure 7: Distribution of Subject Party

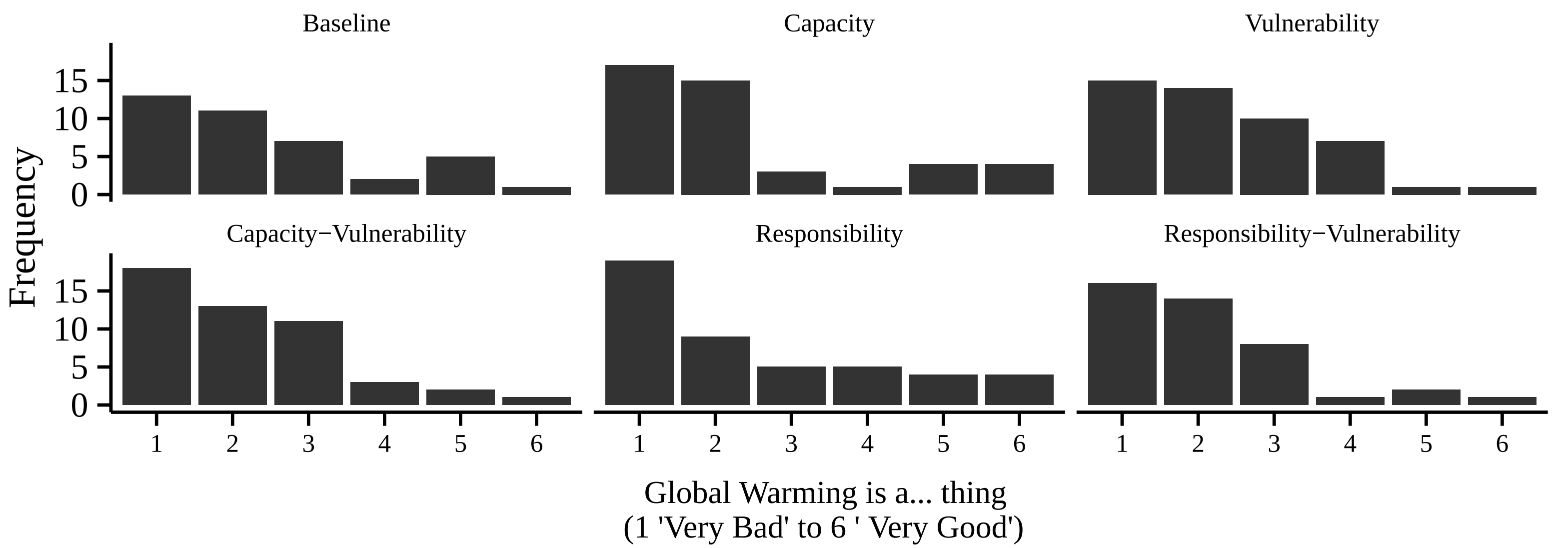


Figure 8: Distribution of Effects of Global Warming

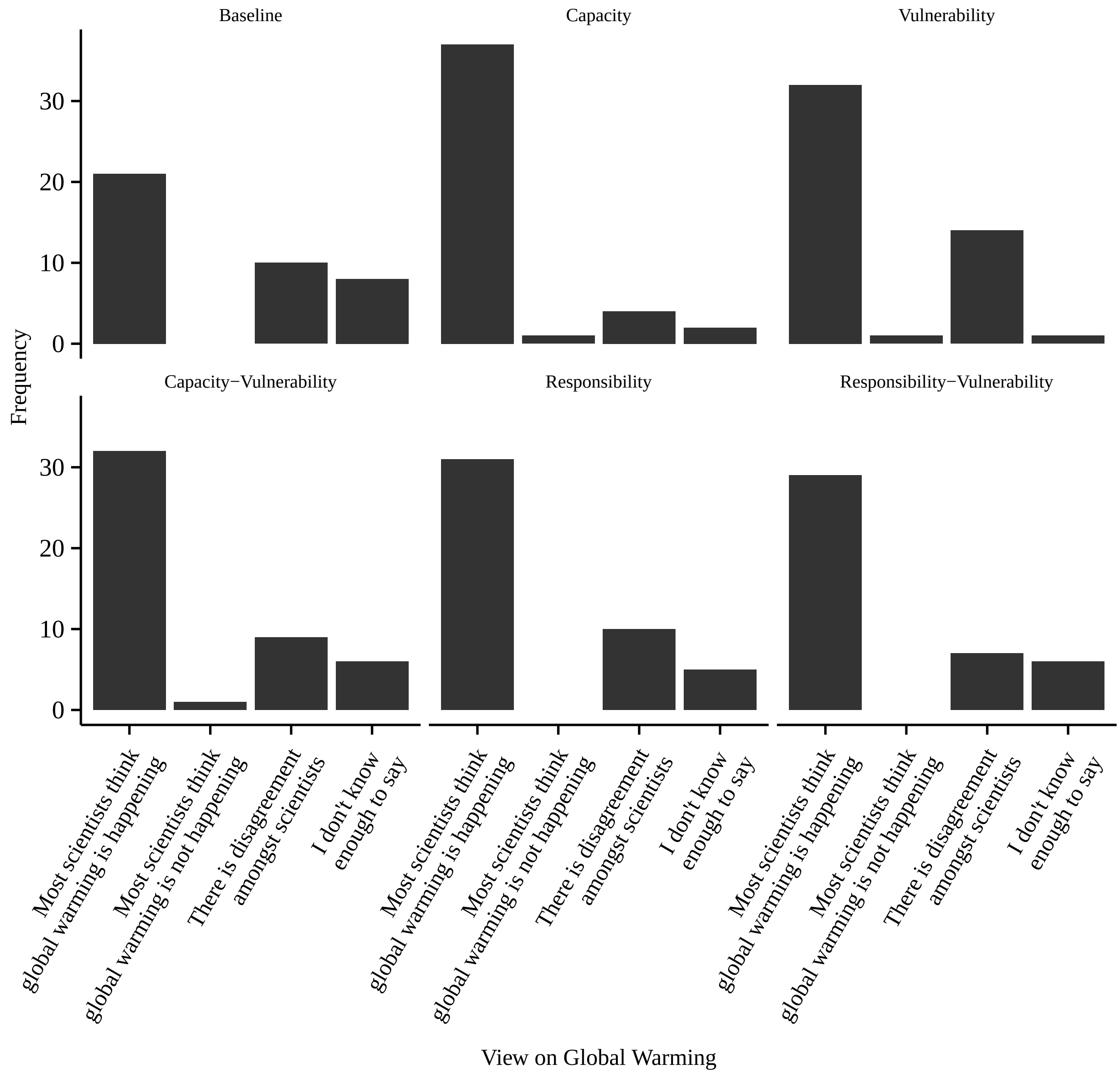


Figure 9: Distribution of Opinion of Evidence of Global Warming

Appendix B Instructions to subjects

Before playing variations of the CBG, subjects received an *overview handout*, which described initial endowments and the set of actions available to each subject, and a *handout supplement*, which included further details, such as the collective risk schedules. The text of the handouts is presented in sections B.1 and B.2. To ensure that participants understood the rules of the CBG, we required them to attend a presentation which reviewed information in the handouts and offered them the opportunity to ask questions. To illustrate this step, sample instructional slides for the *CP-VN* treatment are provided in section B.3.

B.1 Overview handout

Overview

- We will randomly pair you with another individual and then assign you to one of two positions: Person A or a Person B.
- You will then be paired with another person. We call this a "pairing."
- After the pairing ends, you will be paired randomly with another person, and randomly assigned to be either Person A or Person B.
- Overall, you will be paired anonymously with multiple individuals in today's session. You will be paid points based on one randomly chosen pairing. Each $1 is worth 6 points.

Pairing overview

- For BL and VN only: At the beginning of each pairing, you will be given 100 points.
- For CP and CP-VN only: At the beginning of each pairing, if you are person A, you will be given 150 points, if you are person B you will be given 50 points.
- For RS and RS-VN only: At the beginning of each pairing, you each will choose between 0 and 100 points to take, and keep.
- For RS and RS-VN only: You and the other person you are paired with must decide how to split a cost. This cost equals the sum of what you and the other person took in the beginning, divided by 2.
- For BL and VN only: You and the other person you are paired with must decide how to split a cost that is subtracted from the 100 points that you begin with.
- For CP and CP-VN only: You and the other person you are paired with must decide how to split a cost that is subtracted from the 150 (if Person A) or 50 (if Person B) points that you begin with.

- For RS and RS-VN only: You and the other person you are paired with must decide how to split a cost. This cost equals the sum of what you and the other person took in the beginning, divided by 2.

- In the first round, Person A starts off as the "offerer" and the Person B starts off as the "responder." The offerer proposes a division of the costs. The responder can accept or reject this division.
- If accepted, both people earn their starting points minus the cost they agree to and the pairing ends.
- If rejected, there is a possibility that your pairing will end, and you will no longer be matched with that person. We will call this the "end state". It occurs with a probability called the "end probability" which the computer uses to randomly decide when the end state is reached.
- If the "end state" is reached, the computer decides for each person whether they keep your points from the round or you lose them. We call this the "loss probability".
- If the pairing does not end, you and the other person you are paired with play another round, with the roles switching. If you began as the offerer you will now be the responder, and vice versa.
- The pairing ends when (i) both you and the person you are paired with accept some division; or (ii) you reach the "end state".
- If an agreement is not reached in round 8, the pairing will automatically move into the end state.
- After you finish this pairing, you will randomly be paired with another person. At no point will you know the identity of the person you are paired with.
- Out of the 8 times you paired off, **one random pairing will be selected and you will be paid 1 dollar for every 6 points you earned during that pairing.**

What else should I know?

- For RS and RS-VN only: During a pairing it is important to keep track of how much cost you are dividing, and the end and loss probabilities.
- For BL, VN, CP, and CP-VN only: During a pairing it is important to keep track of the end and loss probabilities.
- Sometimes, you and the person you are paired with may take less time to complete the pairing than others. In this case, please wait until we re-pair you with someone else.

B.2 Handout supplement

This handout provides more specific details about the maximum number of rounds for a pairing between you and another person, the number of points with which you start off, the cost that you must split, the "end probability", and the "loss probability."

Points with which you begin the pairing

For BL and BL-VN only: Person A and Person B begin with **100 points**.

Person A (start as offerer)	Begin with **100 points**
Person B (start as responder)	Begin with **100 points**

For CP and CP-VN only: Person A begins with **150 points** and Person B begins with **50 points**.

Person A (start as offerer)	Begin with **150 points**
Person B (start as responder)	Begin with **50 points**

For RS and RS-VN only: Person A and Person B each choose between **0 and 100 points**.

Person A (start as offerer)	Begin with **? points**
Person B (start as responder)	Begin with **? points**

Costs that you must split

For BL, VN, CP, and CP-VN only: The cost that the two of you must agree to split is **100 points**.

Cost	**100 points**

For RS and RS-VN only: The cost that the two of you must agree to split is the average of points chosen by A and points chosen by B.

Cost	**(A points + B points)** / 2

End probability

Recall that the "end probability" is the probability that the computer uses to randomly decide when the end state is reached. In this interaction, the end probability remains constant at **25%**.

Following each round in which the second-mover rejects the first-mover's proposed division, the interaction will reach the end state with a **25% probability**.

End probability	**25%**

After **8 rounds**, the interaction will automatically reach the end state.

Loss probability [for BL, CP, and RS only]

If you reach the "end state", the "loss probability" is the probability that the computer uses to randomly decide whether you lose all your points earned in the round. In the first round, the loss probability for Person A and Person B is 20%.

This means that, if Person B rejects the Person A's proposed division and if the game goes into the "end state", there is a 20% chance that Person A will lose all points, and there is 20% chance that Person B will lose all of their points.

The loss probability increases over every round of the interaction. This means that, as rounds continue, each person faces a greater probability of losing all points earned. The loss probabilities are summarized below:

	Round							
Individual	1	2	3	4	5	6	7	8
Person A (starts as "offerer")	20%	30%	40%	50%	60%	70%	80%	90%
Person B (starts as "responder")	20%	30%	40%	50%	60%	70%	80%	90%

Loss probability [for VN, CP-VN, and RS-VN only]

If you reach the "end state", the "loss probability" is the probability that the computer uses to randomly decide whether you lose all your points earned in the round. In the first round, the loss probability for Person A is 5% and for Person B it is 20%.

This means that, if Person B rejects the Person A's proposed division and if the game goes into the "end state", there is a 5% chance that Person A will lose all points, and there is 20% chance that Person B will lose all of their points.

The loss probability increases over every round of the interaction. This means that, as rounds continue, each person faces a greater probability of losing all points earned. The loss probabilities are summarized below:

	Round							
Individual	1	2	3	4	5	6	7	8
Person A (starts as "offerer")	5%	10%	15%	20%	25%	30%	35%	40%
Person B (starts as "responder")	20%	30%	40%	50%	60%	70%	80%	90%

B.3 Sample presentation slides (CP-VN)

Please do not touch the computers

Please silence your cell phones

Please pay attention to the following instructions, as they will describe how the experiment works and how you will be paid based on your decisions

Experiment

- You will engage in **8** pairings. Each time, we will randomly pair you with another subject in the lab.
- In each pairing, you will make decisions that affect how much money you will make from today's experiment.
- All decisions will be made anonymously through a computer interface.

Experiment

- Out of the **8** pairings that you have, your payment will be determined from one randomly selected pairing, so it is important to pay attention throughout the experiment.
- You will be paid privately at the end of the experiment.

Role Assignment

- In the experiment you will be randomly divided into groups of 2
- Within that group, the computer will randomly assign you to be either Person A or Person B.
- Person A begins the pairing as the **offerer** and Person B begins as the **responder**.
- Each pairing involves a series of rounds.
- In each round, A and B switch roles.

Endowments

Person A and Person B each start the pairing with **YA** and **YB** points respectively, and must agree to divide a **cost** of **Q** points between one another.

In the following instructions, we assume that Person A is the offerer and Person B is the responder.

Dividing Costs

1) Person A (the offerer) suggests a division in the cost (**Q**), with Person A bearing cost **C** and Person B (the responder) bearing cost (**Q-C)**.

Accept/Reject

2) Person B either **accepts** or **rejects** the division proposed by the offerer.

   - **If Person B accepts** the division, the pairing ends.
   - Person A ends the pairing with **(YA – C)** points, and Person B ends the pairing with **(YB –Q+ C)** points.

Rejection: End State (1/3)

If **Person B rejects the division** there is a possibility that the computer terminates your pairing. We call this the "end state."

The probability this occurs is the "end probability" or ***e***.

The "end probability" can remain the same or change every round.

Both you and the person you are paired with can see each other's end probabilities.

Rejection: End State (2/3)

<u>If the end state is reached</u>:

The computer will decide whether each person keeps all points or loses all points.

We call the probability of losing all points ***p*** and this probability may change every round.

Person A and B each have their own value of *p* for each round. The values will be provided to you in a table.

Rejection: End State (3/3)

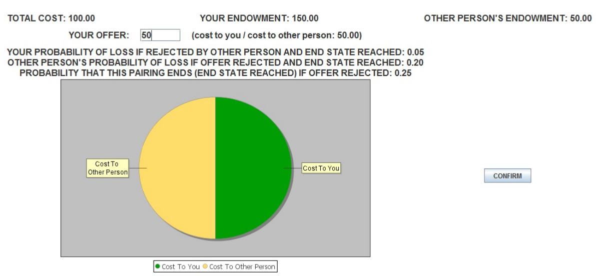
<u>If Person B rejects *and* the pairing does not end</u>:

The steps outlined in the last five slides are repeated, with the roles played by Person A and Person B **<u>reversed</u>—**now Person B becomes the offerer and Person A becomes the responder.

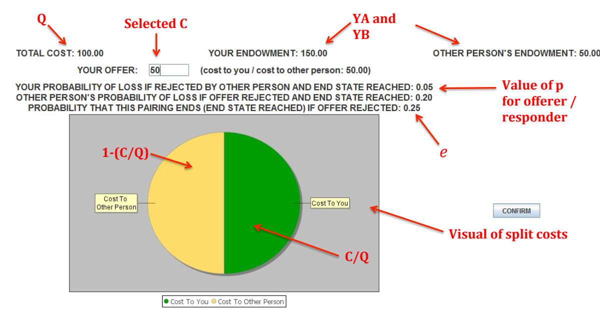
In each subsequent round, the players alternate roles, with A being the offerer in the odd rounds, and B being the offerer in the even rounds.

As long as offers are rejected, the pairing repeats for up to **8 rounds,** after which the "end probability" will go to 100%.

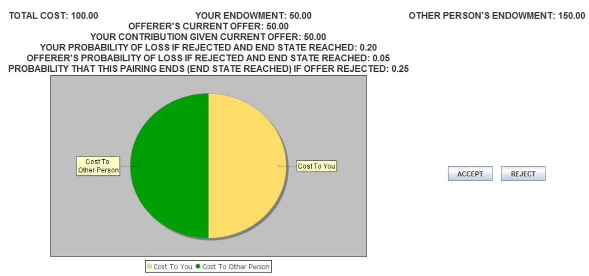
Screenshots of Interface (offerer)



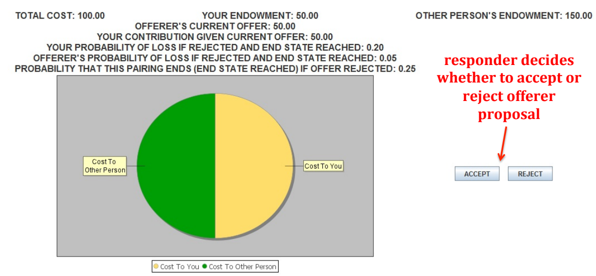
Screenshots of Interface (offerer)



Screenshots of Interface (responder)



Screenshots of Interface (responder)



Re-pairing

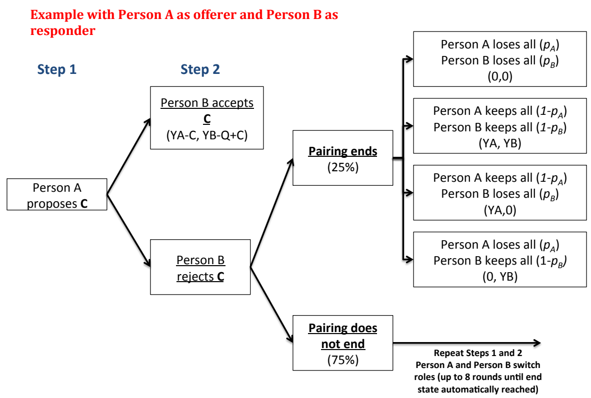
- Today's experiment will involve multiple pairings. You will be randomly paired with another participant each time.
- This re-pairing is anonymous and you will not know the other participant's decisions from previous rounds.

Increasing Probability of Loss

If the responder rejects the offerer's proposed division and the pairing reaches the "end state," the following table displays the probability that each person loses all points increases over every round (*p*).

	Round							
	1	2	3	4	5	6	7	8
Person A	**5%**	**10%**	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**
Person B	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**

Decision Diagram



Questions

- Q: Will anyone ever associate my name with the choices I make?
  - A: No, everything is anonymous.

- Q: How will I earn points?
  - A: If you and the other participant you are paired with reach an agreement on the division of costs, points are based on the division.
  - Otherwise, points are based on whether the computer randomly determines you lost all points or kept all points.
  - Your final payment for today will be determined from one randomly selected pairing

Questions

- Q: Which of the following probabilities are the same for both the offerer and the responder in one particular round?

  - the end probability, *e*
  - the probability of losing all points, *p*
  - both of the above
  - none of the above

Questions

- Q: Which of the following probabilities are the same for both the offerer and the responder in one particular round?

  - the end probability, *e*
  - the probability of losing all points, *p*
  - both of the above
  - none of the above

Questions

- Q: Which of the following probabilities remain constant for a person across the different rounds?
  - the end probability, *e*
  - the probability of losing all points, *p*
  - both of the above
  - none of the above

Questions

- Q: Which of the following probabilities remain constant for a person across different rounds?
  - the end probability, *e*
  - the probability of losing all points, *p*
  - both of the above
  - none of the above
- ANSWER: the end probability remains constant at 25% until the 8^{th} round, when it increases to 100%. The probability of losing all points increases every round.

Questions

- Q: How much money will I earn?
  - A: After you have **8** pairings, the computer will randomly select one of those pairings. Based on the number of points that you earned in that pairing, you will be paid $1 for every 6 points—so that 100 points is worth approximately $17.

Questions

- Before we begin, do you have any questions?

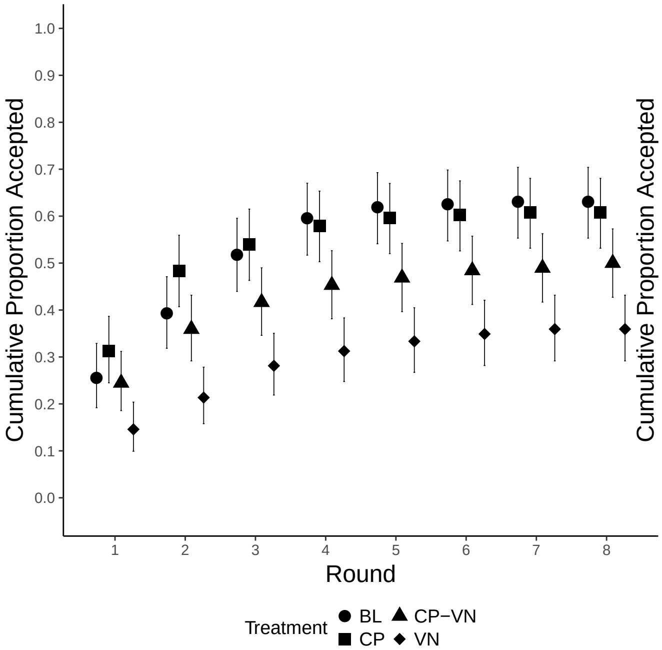
Log IN

- When you are asked to submit your name, please enter your ID code EXACTLY as it appears on your card. Failure to do so will prevent us from paying you.

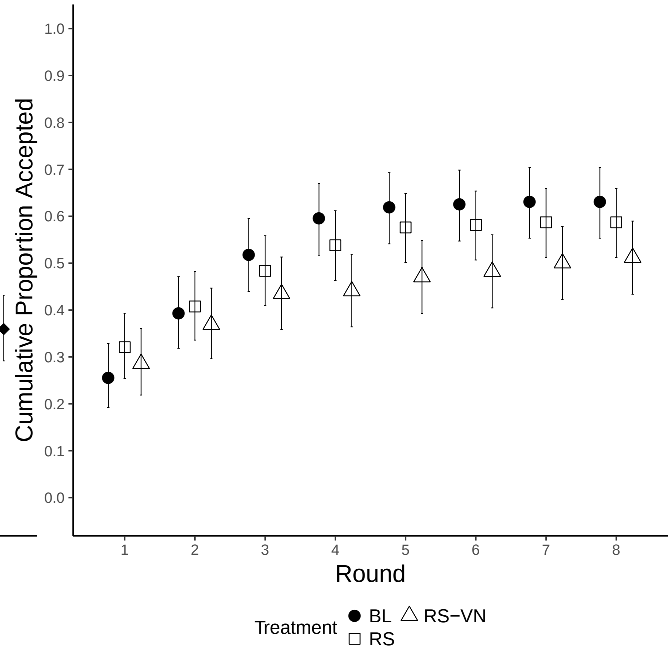
- Please remember there is no talking during the experiment with other participants.

Appendix C Results for full match (eight rounds)

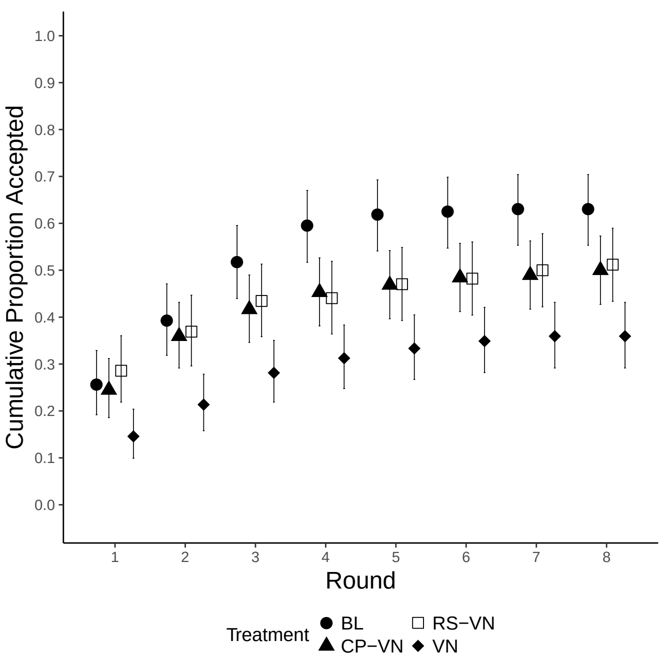
Here we replicate Figures 3 and 4 for all eight rounds.



(a) Exogenous Endowments - Plot of Cumulative Acceptance per Round

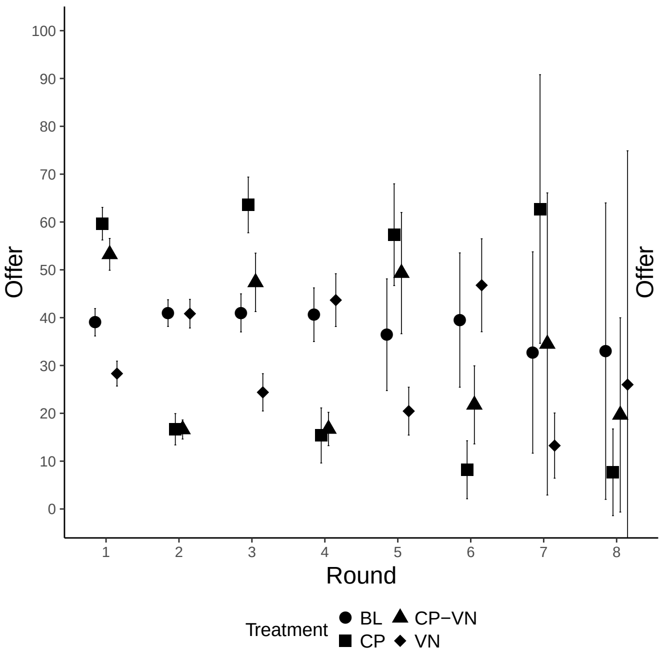


(b) Endogenous Endowments - Plot of Cumulative Acceptance per Round

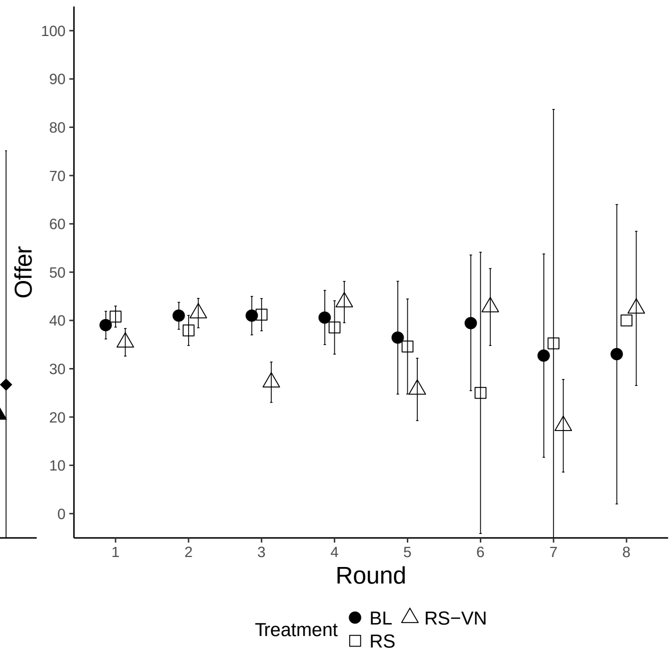


(c) Uneven Probability Treatments - Plot of Cumulative Acceptance per Round

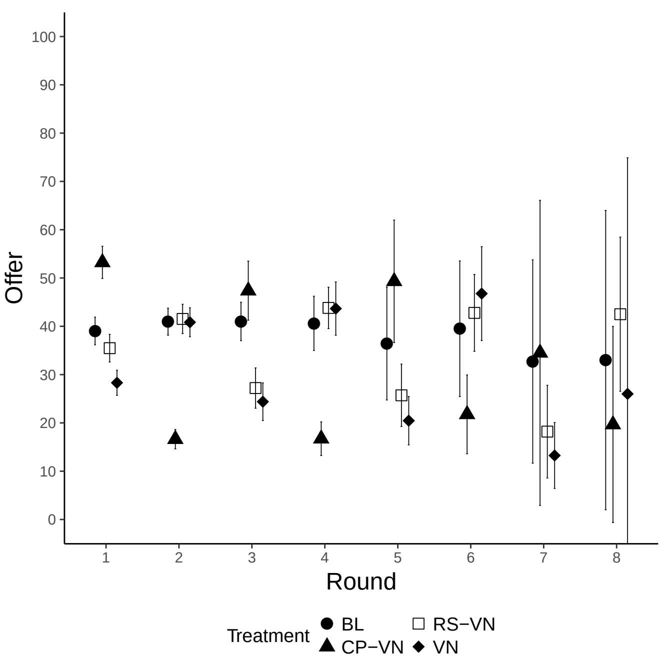
Figure 10: Cumulative Proportion of Acceptances Across Rounds



(a) Exogenous Endowments - Mean Plot of Offers per Round



(b) Endogenous Endowments - Mean Plot of Offers per Round



(c) Uneven Probability Treatments - Mean Plot of Offers per Round

Figure 11: Average Offers Across Rounds

Appendix D Offer/Counter-offer analysis

In this appendix we present several analyses of offer/counter-offer dynamics. Figures 12 and 13 reflect patterns in Player B's reaction to Player A's offers.

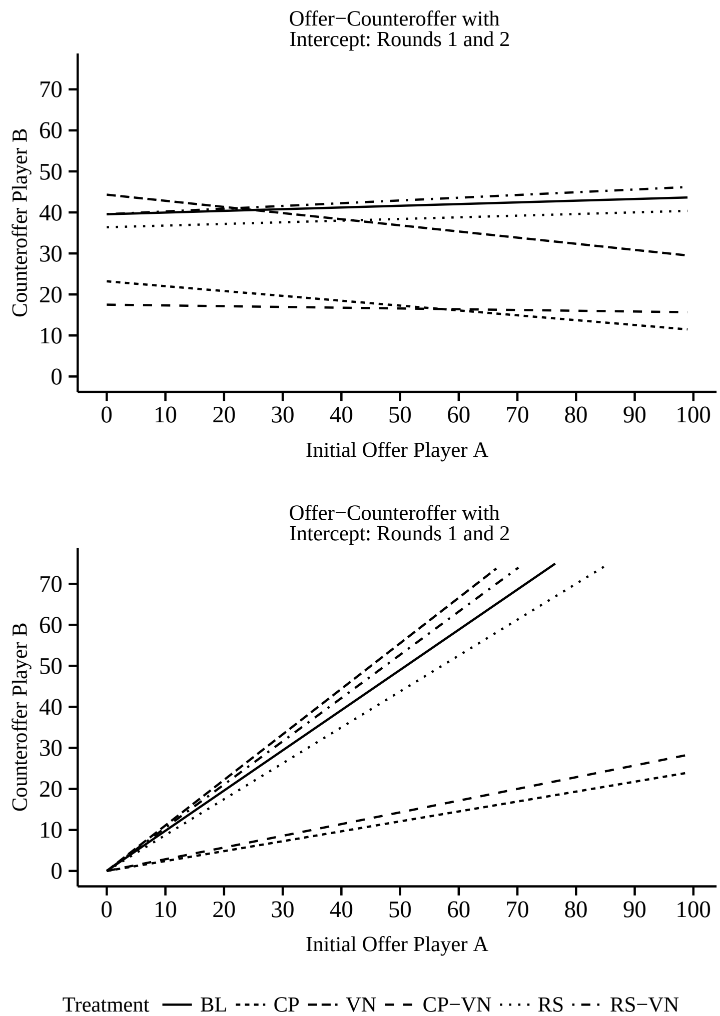


Figure 12: Offer v Counteroffer (Rounds 1 and 2)

The first panel of each figure shows Player B's counteroffers regressed onto Player A's corresponding initial offers with an intercept. We find that Player B counteroffers are consistently lower than Player A offers in round 1 under the *CP* and *CP-VN* treatments, which has been borne out by earlier results. Additionally, given an intercept, Player B's counteroffer decreases in Player A's initial offer, though this may be due to extremely low incidence of Player A offering less than 50.

The second panel removes the intercept, and therefore provides a better idea of each player's sensitivity to the other's offer under various treatments. Player B appears to be most sensitive to Player A's under the *RS-VN* and the *VN* treatments. If Player A offers Player B greater amounts under the *CP* and *CP-VN* treatments, incremental increases in offer comprise a smaller percentage of the total initial offer, and so Player B may be less sensitive to changes. Since Player A tends to give Player B lower offers under the *VN* and *RS-VN* conditions, each incremental increase or decrease in the offer more drastically affects Player B's response. Additionally, examining the *BL* case in the figure, we find that Player B generally adjusts his or her counteroffer to match Player A's initial offer. Although 11a suggests that players were offering one another less than 50, this would likely have been due to Player A's initial offers below 50, in response to which Player B would respond with a similar counteroffer under *BL* conditions.

Figure 13 similarly summarizes Player A's response to Player B's offer in round 2. In an analogous case to Figure 12, we find that Player A is relatively sensitive to changes in Player B's offer under the *CP-VN* and *CP* cases, under which Player B contributes a relatively lower share of the cost. Alternatively, when Player A is expropriating greater offers from a vulnerable Player B under *VN* and *RS-VN*, incremental changes in Player B's offer have a lower effect on Player A's counteroffer.

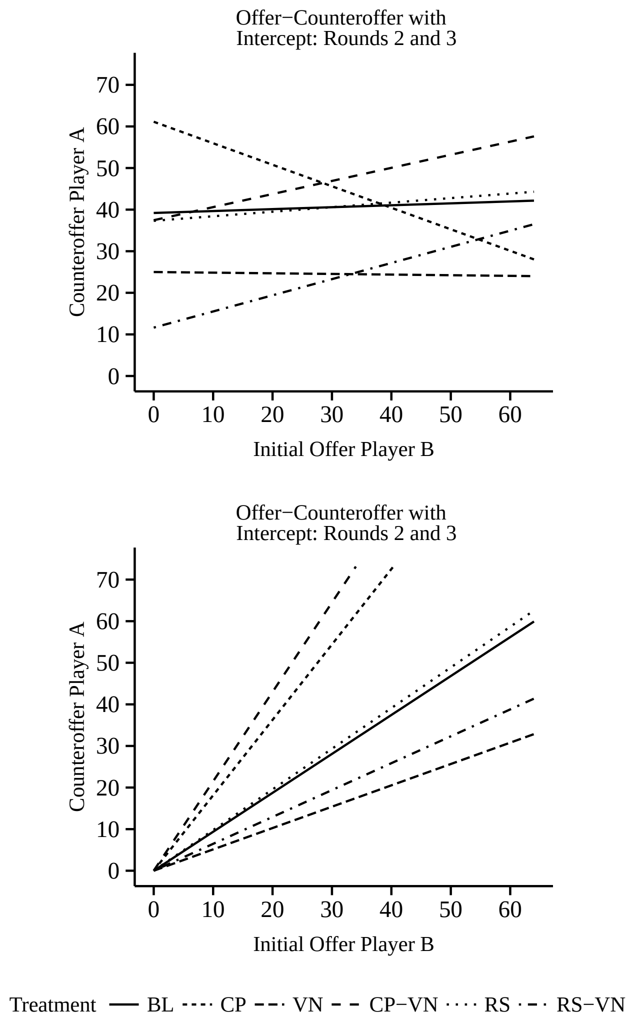


Figure 13: Offer v Counteroffer (Rounds 2 and 3)

As an additional observation, both sets of plots in Figure 13 select cases in which the initial offer was not accepted. Consequently, counteroffers are being made by players who were not willing to accept the initial offer, and so some the plots may be selecting for more demanding players, resulting in overstated slopes in the plots without intercepts, and overstated slopes and/or understated intercepts in plots with intercepts or local regressions.

Appendix E *RS* and *RS-VN* results including asymmetric cases

Our results indicate that continuation values are an important determinant of both the size of the offers and the likelihood of success in the bargaining game. Unlike capacity and vulnerability, the responsibility factor — that is to what extent each party is responsible for the costs of climate change mitigation that must be agreed upon — does not affect continuation values. Nonetheless, given its position as a mainstay of current climate equity discussions and its obvious importance in light of the statements made in many INDCs, we investigate the responsibility factor here in the appendix. When bargaining begins, responsibility is already incorporated into the players' capacities, and thus has no direct effect on continuation values. It should therefore not be taken into consideration by rational actors. As a result, the responsibility treatments have continuation values identical to those in their counterpart treatments that do not vary responsibility. We can therefore use the responsibility manipulations to further test our claims regarding the heightened importance of continuation values in the context of bargaining under increasing collective risk. For *BL* and *VN* we conducted "placebo" conditions of this type with endogenous responsibility for the cost of mitigation. We will refer to these as *RS* and *RS-VN* respectively.

Responsibility treatment

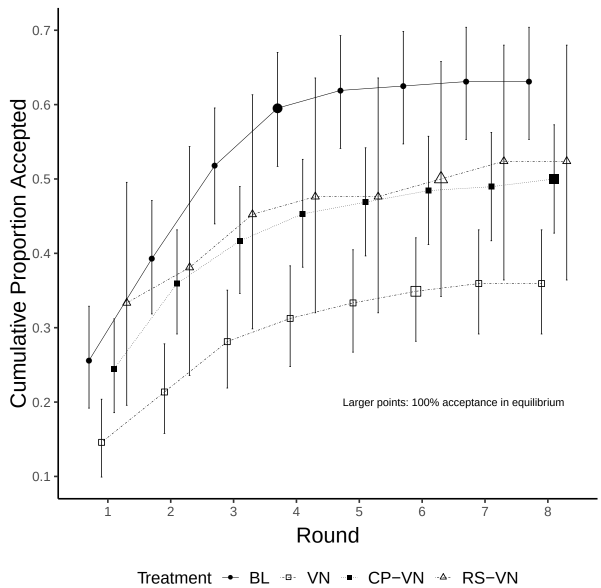
To capture *responsibility* considerations in the climate bargaining framework, we allow each player to endogenously determine his or her initial endowment. The selection of endowments affect total mitigation costs, which amount to the average of the endogenously selected endowments. In these **responsibility** (*RS*) conditions, the subjects, with full knowledge of how the mitigation costs are calculated, will each be allowed to extract an amount from a common pool resource valued at 200. Each player may select an initial endowment between 0 and 100. In all other respects, *RS* conditions are identical to *BL*. This manipulation allows us to compare the results of the *RS* conditions with the *BL* conditions, thereby isolating the role of responsibility and allowing us to investigate the effect of the "polluter pays" principle.

Responsibility-Vulnerability Treatment

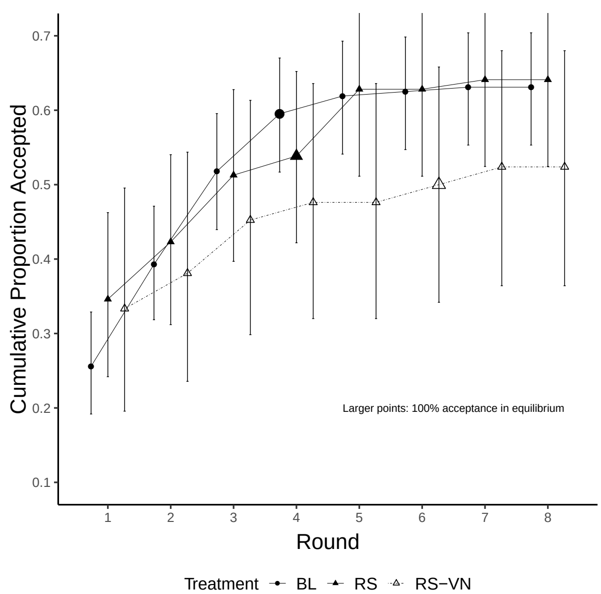
To further understand the effects of *RS*, we combine it with *VN*. The **responsibility-vulnerability** (*RS-VN*) treatment is characterized by endogenously determined initial endowments. Subjects have complete information about extraction amounts and the function determining the cost. combined with the parameterization of each player's collective risk schedule from the *VN* condition, that is (0.05, 0.10, ...0.4) and (0.2, 0.3, ...0.9) respectively. By design, the *RS* and *RS-VN* conditions do not affect continuation values over and above those operating via endowments and the risk schedules (which are identical to those in their counterpart treatments, *BL* and *VN*), so they provide a further test of the importance of continuation values in bargaining under collective risk.

Why do we call these "placebo" conditions? That is because in the instances in which each player choose the maximum possible endowment for themselves – 100 – the continuation values at that point are identical to those in their respective counterpart treatments. For that reason, we restrict our attention to only those cases. Because choosing the maximum amount is a dominant strategy — and a fairly obvious one at that — these constituted 60.6% of all cases for the *RS* treatment and 75.4% of all cases for the *RS-VN* treatment. For the reasons already discussed, then, these are the true placebo cases. When one or the other partner chooses something other than 100, then the continuation values will not be identical, and comparisons are more difficult. We therefore exclusively focus on these placebo cases.

First we replicate Figure 1 for *RS* and *RS-VN* treatments. (Figure 14).



(a) Uneven Probability Treatments - Plot of Cumulative Acceptance per Round



(b) Endogenous Endowments - Plot of Cumulative Acceptance per Round.

Figure 14: Cumulative Proportion of Acceptances Across Rounds Subsetting *RS* and *RS-VN* Pairings

Again we observe high success rates relative to the theoretical predictions, and we see the same variability in success across conditions. In other words, *RS* looks like *BL* and RS-VN looks like *VN*. Continuation values are held constant across the counterpart pairings, which means that observing the same behavior across treatments is consistent with the importance of continuation values. Again we can ask, compared to the theory's predictions, are the offerers too generous? Are the responders too willing to accept low offers? For that we will replicate figures 2, 3, and 4 from the text, again with data from from the responsibility conditions.

As in figure 2 in the text, figure 15 shows that first round offers by Player A in *RS-VN* are similarly generous to those in *VN*, whereas offers in *RS* are not more or less generous compared to equilibrium, similar to those in *BL*. Player B's behavior in the responsibility conditions also mirrors that in their counterpart conditions: less generous offers in *RS-VN* and *VN*, and more-or-less equlibrium offers in *RS* and *BL*. So, again, the addition of responsibility seems to have done little to affect behavior.

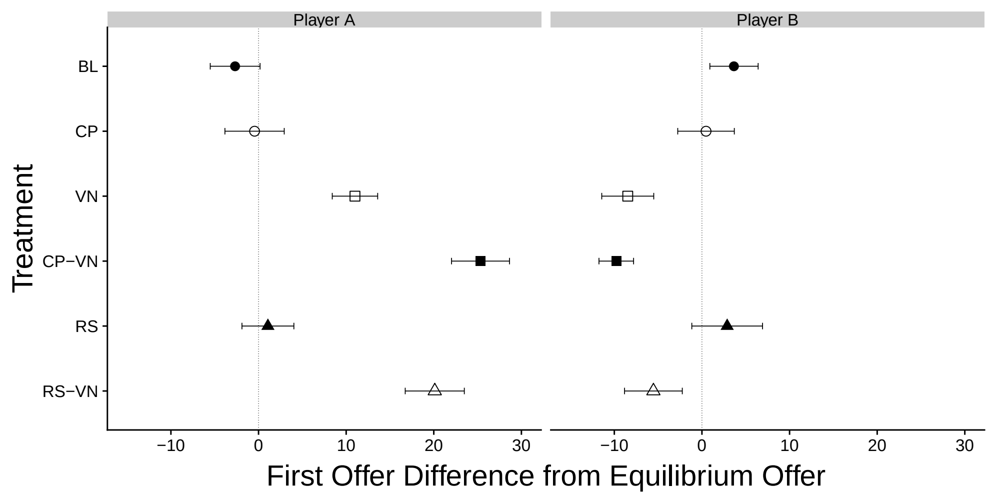
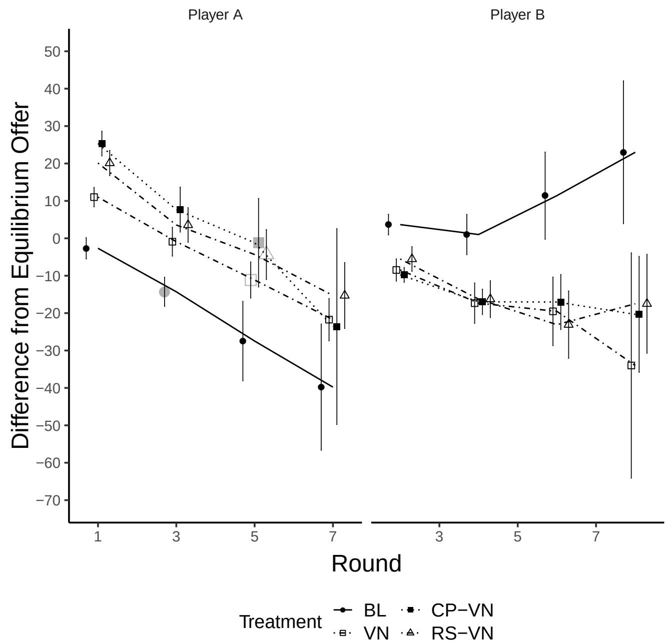


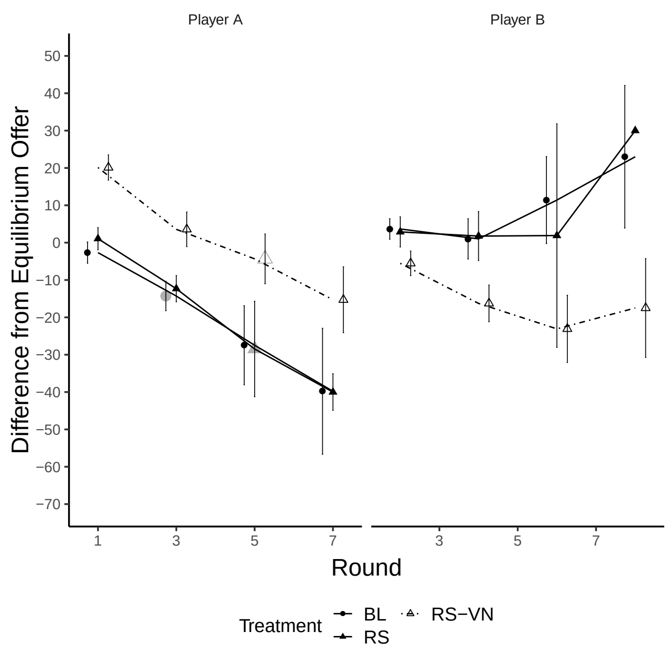
Figure 15: Difference between Observed and Equilibrium First Offers by Type and Condition

Figure 16 displays the trends in the differences presented in Figure 2, but with the addition of the results from *RS* and *RS-VN* over the course of the eight potential rounds. Some interesting patterns emerge. Player A's average offer declines across rounds for the responsibility conditions, as it does for all other conditions. Again, we see the asymmetric vulnerability conditions — including *RS-VN*–group together, with their intercepts shifted upward, indicating more generous offers, relative to equilibrium, than offers in the symmetric vulnerability conditions — *BL*, *CP*, and, now, *RS*. For Player B, the same pattern of behavior also repeats itself. Once again, the addition of responsibility does not seem to have affected behavior in any meaningful way.

In the case of Figure 17, it is somewhat more difficult to visually compare to Figure 4 in the text. Nonetheless, the basic pattern seem to be the same for the responsibility conditions as it is for their counterparts. Once again, responsibility seems not to have had any affect on behavior, and these results — just as in the main text — suggest that responders' low willingness to accept offers — rather than generosity of the offerer — is the major driver of



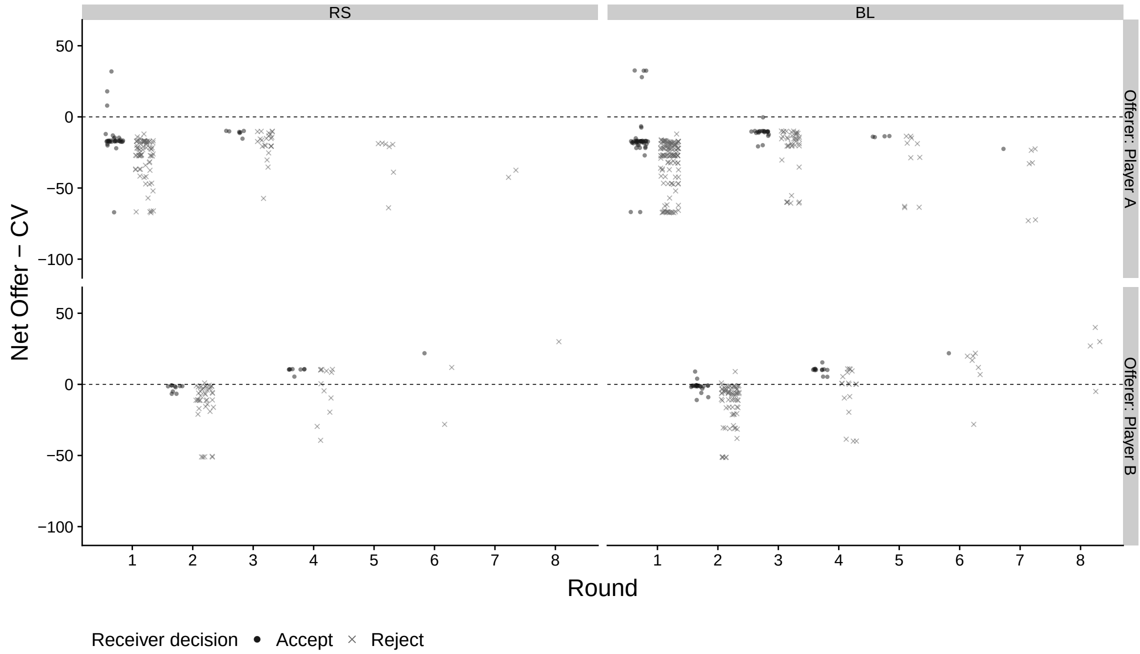
(a) Asymmetric Vulnerability Treatments



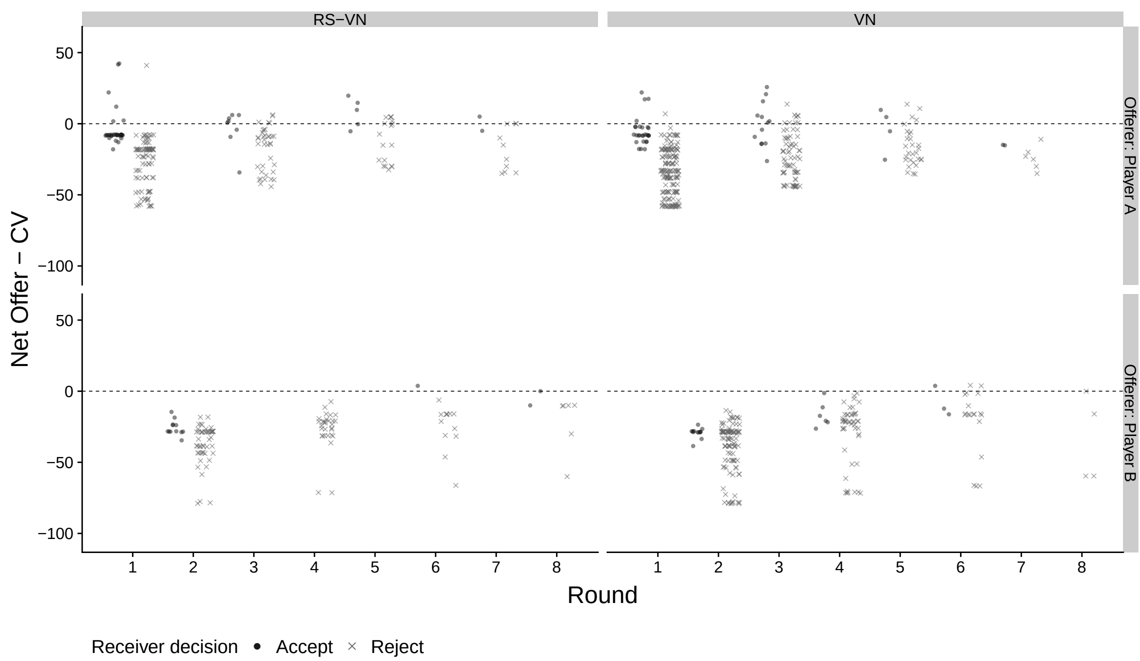
(b) Endogenous Endowments

Figure 16: Round-by-Round Differences between Observed and Equilibrium Offers by Type and Condition

success rates.



(a) Endogenous Treatment Compared to Baseline Treatment



(b) Endogenous Asymmetric Vulnerability Treatment Compared to Asymmetric Vulnerability Treatment

Figure 17: Distribution of Accepted and Rejected Offers by $E_{\text{Responder}} - C_{\text{Responder}} - V_{\text{Responder}}$, Round, Player, and Treatment. The two vertical panels correspond to the *RS* and *RS-VN* treatments, and rounds are distinguished on the horizontal axis, with odd rounds corresponding to cases where Player A is the offerer (on the top panel) and even ones to those where Player B is the offerer (on the bottom panel). The vertical axis describes the difference between net offers ($E_{\text{Responder}} - C_{\text{Responder}}$) and responders' continuation values ($V_{\text{Responder}}$), with the dashed horizontal line indicating cases where the net offer matches the responders' continuation value. Offers accepted by the responder are illustrated by the black points on the left side of each round, and rejected ones by the grey "x"s on the right. Theoretically, we expect that the receiver will accept net offers greater than the continuation value (above the dotted line) at a higher rate than those below the continuation value.