Effects of fairness principles on willingness to pay for climate change mitigation

Journal Article

Author(s):
Anderson, Brilé; Bernauer, Thomas; Balietti, Stefano

Publication date:
2017-06

Permanent link:
https://doi.org/10.3929/ethz-b-000236412

Originally published in:
Climatic Change 142(3), https://doi.org/10.1007/s10584-017-1959-3

Funding acknowledgement:
295456 - Sources of Legitimacy in Global Environmental Governance (EC)

Climatic Change (2017) 142:447–461
DOI 10.1007/s10584-017-1959-3

Effects of fairness principles on willingness to pay for climate change mitigation

Brilé Anderson[1] · **Thomas Bernauer**[1] · **Stefano Balietti**[2,3,4]

Received: 22 September 2016 / Accepted: 21 March 2017 / Published online: 14 April 2017

Abstract Despite the shift from multilateral negotiations on legally binding mitigation commitments to the decentralized nonbinding Intended Nationally Determined Contributions (INDCs) approach in global climate policy, governments and other stakeholders continue to insist that fairness principles guide the overall effort. Key recurring principles in this debate are capacity and historical responsibility. To keep global warming within the internationally agreed 2 °C limit, many countries will have to engage in more ambitious climate policies relative to current INDCs. Public support will be crucial in this respect. We thus explore the implications of different fairness principles for citizens' preferences concerning burden sharing in climate policy. To this end, we implemented an online experiment in which participants (N = 414) played an ultimatum game. Participants were tasked with sharing the costs of climate change mitigation. The aim was to examine how participants' willingness to pay for mitigation was influenced by capacity and historical responsibility considerations. The results show that fairness principles do have a strong effect and that participants applied fairness principles differently depending on their position at the outset. It turns out that participants paid more attention to other players' capacity and historical responsibility when proposing a particular

Electronic supplementary material The online version of this article (doi:10.1007/s10584-017-1959-3) contains supplementary material, which is available to authorized users.

✉ Brilé Anderson
brile.anderson@ir.gess.ethz.ch; http://www.ib.ethz.ch

Thomas Bernauer
http://www.ib.ethz.ch

Stefano Balietti
http://www.networkscienceinstitute.org

1 Center for Comparative and International Studies (CIS), ETH Zürich, Haldeneggsteig 4, 8092 Zürich, Switzerland

2 Network Science Institute, Northeastern University, 177 Huntington Avenue, Boston, MA 02115, USA

3 Harvard Institute for Quantitative Social Science, Cambridge, MA, USA

4 D'Amore-McKim School of Business, Boston, MA, USA

cost allocation and more attention to their own capacity and responsibility when responding to proposals by others. These and other findings suggest that framing climate policy in terms of internationally coordinated unilateral measures is likely to garner more public support than framing climate policy in terms of a global bargaining effort over the mitigation burden.

1 Introduction

In 1992, nearly 200 countries formed the UN Framework Convention on Climate Change (UNFCCC) in order to prevent dangerous anthropogenic interference with the climate. When founding the UNFCCC, states decided that mitigation costs should be distributed based on common but differentiated responsibilities and capabilities. Initial negotiations led to the 1997 Kyoto Protocol, which set modest reduction targets for 37 countries until 2012. Negotiations on a binding follow-up agreement to the Kyoto Protocol failed, largely due to deadlock over how to "fairly" distribute the overall greenhouse gas (GHG) mitigation burden. This impasse was caused by conflict over two potential principles for allocating mitigation obligations and associated costs (Underdal and Wei 2015): historical responsibility (i.e., distributing mitigation costs by countries' contribution to climate change) and capacity (i.e., distributing mitigation costs by countries' financial ability to respond) (Rogelj et al. 2016; Mattoo and Subramanian 2012).

At the Paris COP in December 2015, UNFCCC member states fundamentally restructured the climate governance process. In place of multilateral negotiations over mutually agreed upon legally binding mitigation commitments as in prior years, each country submitted an Intended Nationally Determined Contribution (INDC), which is equivalent to a unilateral nonbinding pledge (Fuglestvedt and Kallbekken 2016). The restructuring of the process intended to overcome the stalemate over which principle to use for distributing the mitigation burden, enable governments to account for differences in domestic circumstances, and avoid a lowest common denominator outcome where the least ambitious country held others back.

Despite the switch to INDCs, the 2015 Paris Agreement failed to resolve the vexing problem of fairness in burden sharing. Governments still pay attention to the behavior of other countries (especially whether others make "fair" contributions) when drafting INDCs. For example, India notes in its INDC that, "even though not a part of the problem, [India] has been an active and constructive participant in the search for solutions." It also states that, "Our objective is to establish an effective, cooperative and equitable global architecture based on climate justice and the principles of Equity and Common But Differentiated Responsibilities and Respective Capabilities, under the UNFCCC".[1] Likewise, Canada states, "With this contribution Canada is affirming our continued commitment to developing an international climate change agreement that is fair, effective and includes meaningful and transparent commitments from all major emitters."[2] Thus, illustrating governments continued emphasis on fairness principles even under the restructured framework of climate governance.

Whether the Paris Agreement will achieve its goal of limiting global warming to 2 °C largely depends on whether citizens will be willing to incur the costs of GHG mitigation over the next few decades. As noted in a recent report by the US Council on Foreign Relations,[3]

[1] UNFCCC, http://www4.unfccc.int/submissions/INDC/Published%20Documents/India/1/INDIA%20INDC%20TO%20UNFCCC.pdf (Accessed 28 November 2016)

[2] UNFCCC, http://www4.unfccc.int/Submissions/INDC/Published%20Documents/Canada/1/INDC%20-%20Canada%20-%20English.pdf (Accessed 28 November 2016)

[3] Council of Foreign Relations, http://blogs.cfr.org/sivaram/2015/12/12/two-cheers-for-the-paris-agreement-on-climate-change/ (Accessed 28 November 2016)

"Climate change policy is almost entirely about domestic policy, and domestic policy is mostly driven by domestic politics."

Recent public opinion polls and survey experiments find robust public support for unilateral GHG mitigation policies (i.e., GHG reductions that are not contingent on whether other countries do the same), even in large GHG emitter countries like India, the USA, and China (Bernauer and Gampfer 2015; Bernauer et al. 2014; Tingley and Tomz 2013). Public support for unilateral climate policy, which is essentially what the Paris Agreement envisages, remains strong even when confronted with information on the high costs of these measures and potential free riding by other countries (Bernauer and Gampfer 2015).

While public backing of unilateral climate policy bodes well for the implementation of current INDCs, present commitments will collectively fail to keep global warming below 2 °C. This means that some of these commitments, notably those of larger emitter countries, will need to become more ambitious.[4] This also implies, in turn, that restructuring the global climate governance process only temporarily ended the gridlock in global climate governance. Fairness in burden sharing has resurfaced in debates over INDCs, and its relevance is likely to increase further as INDCs become more onerous.

In light of this, it is important to understand how the most prominent principles of burden sharing in climate policy (i.e., historical responsibility and capacity) resonate with the public and affect their willingness to pay for GHG mitigation. More specifically, we are interested in whether and to what extent the public cares about principles of historical responsibility and capacity in climate policy and if governments should continue to use these principles as justifications for more ambitious goals.

Existing research on individuals' willingness to pay for GHG mitigation in light of different fairness principles relies primarily on surveys or discrete-choice experiments embedded in surveys (see Bechtel and Scheve 2013; Carlsson et al. 2011; Carlsson et al. 2013; Schleich et al. 2016). The main limitation of this research is that it does not delineate the link between costs of GHG mitigation to the individual and different principles of burden sharing (see Section 2). Ultimately, it captures individuals' "willingness to accept" different fairness principles rather than their willingness to pay for the implications of different fairness principles.

Game experiments are useful in connecting fairness preferences and willingness to pay for GHG mitigation (Almås et al. 2010; Cappelen et al. 2013; Levitt and List 2007).[5] In such experiments, individuals act on their own behalf and receive payoffs based on their own and other players' behavior as well as outcomes in the game (Bernauer and Gampfer 2015). We use an ultimatum game (explained below) and calculate players' threshold values in equilibrium. The threshold is equivalent to the highest values each player is willing to pay. For simplicity, we refer to this as "willingness to pay." To our knowledge, Gampfer (2014) is the only study to date that uses a laboratory experiment to shed light on individuals' fairness preferences in GHG mitigation.

Similar to Gampfer (2014), two players distribute the *costs* of GHG mitigation in our experiment. Each player has a certain level of wealth at the start of the game (representing the

[4] Climate Action Tracker, http://climateactiontracker.org (Accessed 28 November 2016)

[5] Public goods games are used to *simulate* climate negotiations to understand how different factors affect the production of a public good (see Barrett 2003, 2006, 2011; Buchanan et al. 2009; Burton-Chellew et al. 2013; Dutta and Radner 2009; Milinski et al. 2008, 2011; Tavoni et al. 2011). Our interest is in the *behavior of a participant* and how their behavior is motivated by fairness preferences meaning an ultimatum game is more appropriate.

"real world" concept of capacity) and their contribution to the likelihood of a climate catastrophe (representing the "real world" concept of historical responsibility). The first player, referred to as the Proposer, offers to pay a certain amount of the GHG mitigation cost. The Responder decides whether to accept or reject the Proposer's offer. If players end up accepting the costs of mitigation, a climate catastrophe is averted. If not, a climate catastrophe will occur with a certain probability.

We modify the experimental design, analysis, and sample used by Gampfer (2014). We improved the experiment to better explain causal mechanisms, which will be discussed in greater detail in Section 2. In the analysis, we explicitly model the strategic interaction of the ultimatum game and analyze the effects of players' personal characteristics (altruism, willingness to take risks, and ecological concern) on willingness to pay unlike Gampfer (2014) who used an OLS and omitted participants' traits. Prior research shows that individuals deviate from the rationally optimal offer because of altruism and risk aversion (Andreoni and Miller 2002; Bolton and Ockenfels 2000; Fehr and Schmidt 1999; Gintis et al. 2003). We thus control for these factors to make sure that, if individuals offer to pay more than is rationally optimal, this is due to their historical responsibility or capacity and not due to risk aversion or altruism. Lastly, our sample is more diverse than Gampfer (2014) since we recruited participants from Amazon Mechanical Turk (AMT) instead of university students.

The results show that a Proposer's willingness to pay is significantly affected by the Responder's capacity and historical responsibility. This means that if the Responder is wealthy or historically responsible, then the Proposer's willingness to pay for mitigation significantly decreases. We also find that Proposers' willingness to pay for mitigation is moderated by altruism. The more altruistic the Proposer is and the higher capacity or historical responsibility she possesses, the more willing she is to pay for GHG mitigation. In contrast, a Responder is significantly affected by her wealth and historical responsibility, regardless of her altruism, and is unaffected by the Proposer's wealth or historical responsibility. This means that the higher the Responder's wealth and historical responsibility are, the more willing the Responder is to pay for mitigation regardless of how much the Proposer can do or has contributed to the problem. We will discuss the policy implications of this finding in the concluding section.

In the remainder of this paper, we first review the relevant literature and present the design of the experiment. This is followed by a discussion of the results and their policy implications.

2 Burden sharing principles and public opinion

Debates in policy-making and academic circles over how to share the costs of GHG reductions have converged on two principles (Underdal and Wei 2015): historical responsibility and capacity (Rogelj et al. 2016; Mattoo and Subramanian 2012).

Developing countries (e.g., those organized in the G77) prefer to allocate the mitigation burden, and thus mitigation costs, by historical responsibility (also known as the polluter-pays principle) (Underdal and Wei 2015). This would mean that industrialized countries bear the bulk of the mitigation burden since they are responsible for most GHG emissions from the industrial revolution until the 1980s. Since robust scientific evidence regarding climate change only became available from the 1990s onwards, these countries assert that it is unfair to hold them accountable for past emissions. In contrast, developed countries prefer to distribute mitigation costs along the lines of countries' capacity, which is also known as the ability-to-pay principle and is usually conceptualized in terms of a country's average income level (e.g., GDP per capita). This means that any country that has the financial capacity to reduce GHG emissions should do so. It also implies that emerging economies, such as China and India, reduce GHG emissions. Yet, China and India defend their right to develop uninhibited from GHG mitigation obligations as industrialized countries did in the past. Thus, translating either principle into a feasible burden sharing framework is challenging.

As noted above, public opinion acts as an important constraint on and facilitator of climate policy. Cost and fairness preferences, in turn, are key elements in opinion formation (Bechtel and Scheve 2013, Gampfer 2014). Surveys and survey experiments have been the prevailing method of investigating the public's preferences for distributing GHG mitigation costs (e.g., Bechtel and Scheve 2013; Carlsson et al. 2011; Carlsson et al. 2013; Schleich et al. 2016). The main limitation of this work is that the link between fairness principles and costs to the individual are often underspecified.

Schleich et al. (2016) use a survey to investigate individuals' fairness preferences in mitigation with respect to different burden sharing rules amongst representative samples from China, Germany, and the USA. They found survey participants ranked distributive justice principles in the same way across countries.· However, the costs for the country or individual were never explicitly stated. Carlsson et al. (2011, 2013) and Bechtel and Scheve (2013) use discrete-choice experiments embedded in a survey. In such an experiment, participants are shown pairs of policies that differ along a set of attributes. For each policy pair, participants select their preferred policy. In all three studies, costs to the average household were treated as a separate attribute independent from the burden sharing principle (see Supplementary Information SI-1 Table 1.1 for details). It is quite possible that individuals' preferences for mitigation would change if the link between preferences and costs to the individual were clearly delineated.

The only experimental study, to our knowledge, that explicitly relates fairness principles and willingness to pay for GHG mitigation is Gampfer (2014). He used an ultimatum game similar to our experiment with university students in a decision laboratory. Players distributed the costs of climate change mitigation as described above. The costs were set by the researcher and demarcated in experimental currency units (ECU). Gampfer analyzed the effects of capacity and historical responsibility on the offer of the Proposer. The results show that the Proposer's offer was influenced by historical responsibility and capacity. With increasing historical responsibility and capacity, Proposers tended to offer to pay a larger share of the overall GHG mitigation cost.

Our experiment builds on Gampfer's work and seeks to address several of its limitations. In terms of experimental design, in Gampfer's (2014) game, players' wealth (i.e., capacity) and historical responsibility were endogenously determined, meaning they were a direct result of players' prior actions within the experiment. That is, players knowingly accumulated their wealth and historical contribution to a climate catastrophe before entering into the ultimatum game. In reality, however, individuals have little control over the capacity and historical responsibility of their country. The positive effect of capacity and historical responsibility on the size of the Proposer's offer may thus be stronger in Gampfer's experiment compared to a game setup where capacity and responsibility are exogenously given. We will address this limitation by implementing the game under conditions of endogenously and exogenously determined capacity and historical responsibility.

Another limitation of Gampfer (2014) is his statistical approach (OLS regression), which uses only the Proposer's offer as the dependent variable. The Responder's decision to accept or reject the Proposer's offer is excluded. This has two implications. First, it is unclear how the attributes of the Responder affect the Proposer's willingness to pay. In light of heated debates over what a fair contribution of industrialized countries to climate change mitigation should be relative to poorer countries, this limitation is obviously important. Second, the statistical model in Gampfer's analysis (OLS) fails to reflect players' strategic interaction and may lead to biased estimates (Signorino and Yilmaz 2003). We will use a model specification and statistical procedure that accounts for the strategic interaction as well as the Responder's willingness to pay, which Gampfer (2014) did not.

As mentioned, we control for personal characteristics (i.e., altruism and risk aversion) to make sure that, if individuals offer to pay more than is rationally optimal, this is due to their historical responsibility or capacity and not due to risk aversion or altruism. The same holds for ecologically concerned individuals who are likely to pay a higher share in order to ensure the costs are covered and acute climate change is avoided. There could also be an interaction between an individual's altruism and their capacity as well as historical responsibility. The more altruistic an individual is, the more weight their capacity and historical responsibility could have on their decisions.

Lastly, our experiment was implemented with a larger and more heterogeneous sample and was carried out online with nodeGame[6] (Balietti 2016), rather than in the lab. The participants were recruited from AMT, an online crowd-sourcing platform.[7]

3 Study design

In this section, we present the experimental design and describe the main variables in the analysis. Figure 1 illustrates the experimental design. We pretested the experiment twice with researchers at the Centre for International Studies (ETH Zürich) and conducted a pilot on 4 February 2015 on AMT.

We recruited participants via AMT on 17 April 2015, 24 April 2015, 30 June 2015, and 9 July 2015 (point 1 in Fig. 1). Several studies show that the quality of data obtained from participants recruited via AMT is as good as the quality of data obtained via other sampling approaches, including those of standard laboratories where participants are physically present (Barabas and Jerit 2010; Paolacci et al. 2010). The norm in lab experiments is to use university students. However, students are unrepresentative of the behavior of the overall population. They tend to have greater cognitive skills, malleable attitudes, and more compliant behavior (Sears 1986; Wintre et al. 2001). While our sample possesses greater diversity than a typical lab experiment, it is still younger, poorer, more male, liberal, and highly educated than the American population (see SI-5 Table 5-2 for details). We control for these factors in the analysis. Descriptive statistics are shown in Table 1.

After recruitment on AMT, participants were automatically directed to nodeGame (point 2 in Fig. 1) (Balietti 2016). NodeGame is an open-source software that enables researchers to program interactive experiments to be played online. After arrival in the nodeGame environment, participants entered a waiting room. From there, groups of four were formed. Participants in the waiting room had to wait until they could be grouped with three others (see point 3 in Fig. 1) or 10 min had passed. If grouping failed after 10 min, the respective participant(s) were sent back to AMT for a payout of 1.00 USD. After a group of four formed,

[6] nodeGame, http://nodegame.org (Accessed 1 December 2016)
[7] Amazon Mechanical Turk, https://www.mturk.com/mturk/welcome (Accessed 1 December)

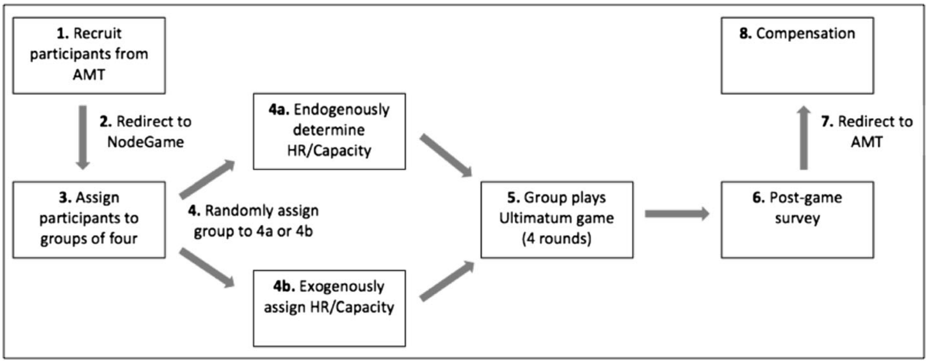

Fig. 1 Experimental design. AMT means Amazon Mechanical Turk, which is an online crowd-sourcing platform. HR means historical responsibility

players read a brief introduction about climate change and the cost (i.e., 30 ECU in the game) to reduce it (see SI-2.1 for the general introduction).

After this introduction, groups were randomly assigned to conditions 4a or 4b (see Fig. 1), which represent two different ways that participants acquired their historical responsibility and capacity before playing the ultimatum game (see point 4 in Fig. 1). In condition 4a, historical responsibility and capacity were endogenously acquired. This means that participants chose the economic growth and GHG emissions of their one-person economy before entering the ultimatum game. Participants selected the economic growth (i.e., low, medium, or high) five times. The exact increase varied: low (increase between 1 to 10 ECU), medium (5 to 15 ECU), or high (10 to 20 ECU). Each level of economic growth corresponded with a contribution to climate risk (i.e., the probability of a climate catastrophe in the ultimatum game): low (0%), medium (2.5%), and high (5%) (see SI-2.3 for exact instructions). All players started with 25 ECU. After the five rounds, their wealth could range from 30 ECU to 125 ECU, and their contribution to climate risk could range from 0% to 25%. The randomization of wealth was meant to ensure that participants with the same level of wealth could incur varying levels of historical responsibility.

In condition 4b, historical responsibility and capacity were exogenously determined, meaning that we randomly assigned participants a given wealth and contribution to climate risk. Therefore, groups assigned to condition 4b moved directly to instructions for the ultimatum game. The relationship between wealth and contribution to climate risk in the exogenous group matched the endogenous group (see SI-2.2). Historical responsibility and capacity were randomly assigned to 213 participants (i.e., exogenous), which resulted in 260 rounds in the analysis. Two hundred one participants were in the endogenous setup resulting in 299 rounds for a total of 559 rounds in the analysis.[8]

After assignment of historical responsibility and capacity, all participants played four rounds of the ultimatum game (see point 5 in Fig. 1). The first round was practice and its

[8] If all players participated in all rounds, we would have 1242 rounds. This is due to a technical malfunction or players voluntarily disconnecting before the end of the game. This is one of the first applications of nodeGame. On the 17 and 24 April, the server overloaded meaning participants were kicked out of the game prior to completing all the rounds. We addressed this problem later on. Other participants left the game prior to completion either voluntarily or their computer malfunctioned. We tested if samples were balanced in the endogenous and exogenous conditions with respect to historical responsibility, capacity, and other player characteristics. We found no significant differences so dropouts should not affect our results.

Table 1 Descriptive statistics

	Unit	Condition[a]	Min	Max	Mean	SD
Altruism	Polar coordinates	*Exogenous*	−16.07	61.39	26.45	13.68
		Endogenous	−7.49	61.39	24.31	13.59
Ecological concern	Continuous scale [0.3]	*Exogenous*	0	1.6	0.95	0.28
		Endogenous	0.27	1.8	0.99	0.26
Risk	0: very risk averse	*Exogenous*	0	10	4.24	2.34
	10: fully prepared to take risks	*Endogenous*			4.10	2.37
Income (annual household)	0: 0 to 10,000 USD	*Exogenous*	0	11	5.00[b]	3.24
	1: 10 to 20,000	*Endogenous*			5.00[b]	3.59
	2: 20 to 30,000					
	3: 40 to 50,000					
	4: 50 to 60,000					
	5: 60 to 70,000					
	6: 70 to 80,000					
	7: 80 to 90,000					
	8: 90 to 100,000					
	9: 100 to 125,000					
	10: 125 to 150,000					
	11: above 150,000 USD					
	12: prefer not to say					
Age	Years	*Exogenous*	18	70	33.43	10.14
		Endogenous	19	66	33.94	10.35
Education	1: no high school	*Exogenous*	1	6	3[b]	1.34
	2: high school	*Endogenous*			4[b]	1.26
	3: some college					
	4: 2-year college					
	5: 4-year college					
	6: postgraduate					
Political ideology	1: left	*Exogenous*	1	5	2[b]	1.45
	2: mostly left	*Endogenous*			2[b]	1.39
	3: center					
	4: mostly right					
	5: right					
Participation	1: participated in an experiment like this before	*Exogenous*	0	1	1[b]	NA
	0: never participated	*Endogenous*			1[b]	NA
Sex	1: male	*Exogenous*	0	1	1[b]	NA
	0: female	*Endogenous*			1[b]	NA

[a] In the endogenous condition, players' historical responsibility and capacity were a direct result of players' prior actions within the experiment. That is, players knowingly accumulated their wealth and historical contribution to a climate catastrophe before entering into the ultimatum game. Players in the exogenous condition were randomly assigned their wealth and historical responsibility

[b] Represents median instead of mean. We did not calculate the standard deviation of participation and sex because they are dummy variables. Forty percent of the endogenous group is female and 60% is male. Likewise, 41% of the endogenous treatment is male and 59% is female. Seventy percent of the endogenous group participated in an experiment like this before, while 65% of the exogenous group previously participated in such an experiment

results were excluded from the analysis. For the three rounds included in the analysis, participants were matched with a different participant each time (see screenshots of the introduction to the game, practice round, and experiment rounds in SI-3).

Participants were compensated based on their remaining wealth at the end of a randomly chosen round of the ultimatum game. Randomly selecting a round for compensation is a common practice in such experiments. It incentivizes participants to behave, in terms of cost benefit calculations, as they would in reality. For the payoff, the conversion rate was 50 ECU for 1 USD plus an initial show up fee of 1 USD. The total payout ranged from 1 USD to 6 USD.

At the end of the game, all participants took a survey measuring participants' altruism, risk perception, ecological concern, and whether they had participated in a game experiment before, followed by a series of sociodemographic questions (see SI-4 for survey questions) (see point 6 in Fig. 1).

We measured risk aversion using: *On a scale of 0 to 10, are you generally willing to take risks, or do you try to avoid taking risks?* (0 = "*very risk averse*"). For altruism, we used the Social Value Orientation measure (Murphy and Ackermann 2013). Participants allocated resources between themselves and another individual six times with differing amounts each time (see SI-4.1). We randomly selected one of these choices as a bonus payout to the respective participant and another participant with the conversion rate of 2 points equal to 1 cent. This setup ensures that participants understood the material consequences of their choices for themselves and another participant. Low values indicate low levels of altruism. To measure participants' ecological concern, we used the New Ecological Paradigm (see Dunlap et al. 2000). Participants were asked whether they agree or disagree with 15 statements (see SI-4.3). The order of the statements was randomized. Following Dunlap et al. (2000), responses were aggregated using the sum and dividing by the total number of questions. Possible scores range from 0 to 3 where low values indicate low levels of ecological concern. After completing the questionnaire, participants returned to AMT where they were compensated for participation in the experiment (see points 7 and 8 in Fig. 1).

The unit of analysis resulting from this experimental design is a round of the ultimatum game. We use an estimator developed by Ramsay and Signorino (2009) to calculate the Proposer and the Responder's willingness to pay using the "games" package in R. The "ultimatum estimator" estimates the effects of different covariates on the Proposer and Responder's willingness to pay simultaneously. The key advantage of the ultimatum estimator, compared to general linear models, such as OLS, is that it models the strategic interaction by using two dependent variables: (1) how much the Proposer would like the Responder to pay (which is used because of the design of the estimator) and (2) the Responder's decision (i.e., whether she/he accepted or rejected the Proposer's offer). The main explanatory variables of interest are the Proposer and Responder's historical responsibility (contribution to climate risk in the game), the Proposer and Responder's capacity (wealth in the game), and whether it matters if this is endogenously determined (i.e., direct result of participants' actions) or exogenously determined (i.e., randomly assigned to participants).

A potential problem in any experiment is experimenter demand effects (EDE). EDE occurs when researchers cue "appropriate" behavior either advertently or inadvertently causing behavioral changes in participants (Zizzo 2010). Since we recruited participants via AMT and redirected them to nodeGame, we can exclude social EDE from the presence of the researcher or as a result social interactions in a lab environment (Zizzo 2010). It is without doubt that cognitive EDE occurred meaning the setup of the game cued participants to our objectives (Zizzo 2010). It is clear that our interest is covering the costs of climate change in the experiment. However, this parallels the real world creating external validity. We wanted to use a contextually rich frame so that participants clearly understood the stakes.

4 Results

Table 5-1 in SI-5 lists correlations between variables when historical responsibility and capacity are exogenously assigned or endogenously determined. We would expect that individuals' risk aversion, altruism, and ecological concern would affect players' behavior when actively deciding on their own economic growth (i.e., the endogenous condition). As we would expect, willingness to take risks is positively and significantly correlated with participants' historical responsibility in the endogenously determined condition. Counterintuitively, ecological concern is positively correlated with capacity and historical responsibility in the endogenous condition. We suspect that individuals feel guilty and, therefore, have greater ecological concern with high levels of capacity and historical responsibility. This could explain why altruism and ecological concern are negatively correlated in the endogenous condition. Altruistic individuals might have opted for lower economic growth to cause less harm to the environment and thus feel less concern for the environment. In both conditions, historical responsibility and capacity are almost perfectly correlated. This reflects reality and avoids participants with low wealth possessing high historical responsibility.

Table 2 displays the main results of the experiment. We tested for multicollinearity, using the value inflation factor,[9] and found evidence of multicollinearity when historical responsibility and capacity were simultaneously included in the model. Therefore, we include these terms in the model separately. Models 1 and 2 in Table 2 include players' capacity, while models 2 and 4 include players' historical responsibility. P represents the Proposer and R represents the Responder. All models are estimated with round fixed effects (i.e., round 2, 3, or 4 of the ultimatum game excluding the first since it was practice).

The results show that a Proposer's willingness to pay for GHG mitigation is not significantly affected by her capacity or historical emissions, as evident from Models 1 and 2, respectively. However, a Proposer's willingness to pay is affected by the Responder's capacity and historical responsibility. That is, if the Responder is wealthy (i.e., has high capacity) or historically responsible, then the Proposer's willingness to pay for the costs of climate change mitigation significantly decreases. Thus, in the role of Proposer, individuals seem to selectively apply fairness principles, namely to the Responder, and not to themselves. However, this only holds the more selfish the individual is. Models 3 and 4 include interaction terms between altruism and capacity as well as altruism and historical responsibility, respectively. We find a significant effect meaning that the more altruistic individuals are, the more they apply fairness principles to themselves as well as the Responder.

We also tested for interaction effects between Proposers' capacity and historical responsibility and their risk aversion as well as ecological concern. However, these effects are insignificant and are thus omitted from the table. The above results remain unchanged when including demographic variables (i.e., education, income, political ideology, age, and gender) (see SI-6 for results). We also controlled for the time when participants started the survey (Huff and Tingley 2015). This effect is significant but the size is less than 0.001.

In contrast, a Responder's behavior is significantly affected by her wealth and historical responsibility and not by the Proposer's wealth and historical responsibility. Thus, as a Responder, individuals appear to apply fairness principles to themselves irrespective of their altruism. The higher the Responder's wealth and historical responsibility, the more willing she is to pay for the costs of climate change mitigation regardless of how much the Proposer can do (capacity) or has contributed to the climate problem (historical responsibility). We find no interaction effects

[9] We calculated the VIF for historical responsibility against capacity, risk aversion, altruism, ecological concern, round, and prior participation. We repeated this for capacity again historical responsibility and the other covariates. Both values were over 5.

Table 2 Results

	Model 1	Model 2	Model 3	Model 4
Proposer				
Endogenous	20.76 (26.32)	7.95 (23.04)	41.20 (27.60)	8.68 (20.11)
P capacity	0.27 (0.35)	–	0.13 (0.42)	–
R capacity	−1.28*** (0.38)	–	−1.44*** (0.34)	–
P historical responsibility	–	0.65 (1.03)	–	−0.26 (0.95)
R historical responsibility	–	−5.38*** (1.53)	–	−5.00*** (1.34)
P altruism	0.33*** (0.09)	0.36*** (0.09)	−0.19 (0.29)	0.04 (0.15)
P ecological concern	−0.71 (2.88)	−1.04 (3.03)	−0.37 (2.99)	−0.31 (3.00)
P risk	−1.37** (0.47)	−1.48** (0.49)	−1.38** (0.47)	−1.43** (0.48)
P participation	−0.19 (2.11)	−0.73 (2.20)	0.25 (2.20)	−0.21 (2.19)
Round 3	18.10 (27.69)	10.85 (26.50)	25.28 (30.20)	14.16 (23.57)
Round 4	−14.68 (30.04)	−7.18 (26.90)	−24.13 (32.17)	−4.78 (23.47)
P altruism × P capacity	–	–	0.01* (0.00)	–
P altruism × P HR	–	–	–	0.02* (0.00)
Responder				
Endogenous	6.50 (7.95)	2.65 (6.61)	11.91 (7.83)	2.78 (5.80)
P capacity	−0.09 (0.10)	–	−0.08 (0.11)	–
R capacity	−0.28** (0.11)	–	−0.33** (0.10)	–
P historical responsibility	–	−0.23 (0.29)	–	−0.32 (0.25)
R historical responsibility	–	−1.26*** (0.42)	–	−1.17*** (0.36)
R altruism	−0.01 (0.02)	0.02 (0.02)	−0.09 (0.08)	0.02 (0.05)
R ecological concern	1.19 (0.86)	1.25 (0.86)	1.32 (0.83)	1.35 (0.85)
R risk	0.18 (0.12)	0.22 * (0.12)	0.20* (0.12)	0.23** (0.12)
R participation	0.44 (0.62)	0.71 (0.63)	0.42 (0.61)	0.72 (0.62)
Round 3	4.84 (8.43)	2.38 (7.60)	6.82 (0.00)	3.39 (6.81)
Round 4	−5.44 (8.85)	−2.96 (7.60)	−76.12 (8.75)	−2.28 (6.66)
R altruism × R capacity	–	–	0.00 (0.00)	–
R altruism × R HR	–	–	–	0.00 (0.00)
Log-likelihood	−1587.681	−1589.196	−2142.23	−2141.55
N (number of rounds)	559	559	559	559

Participation controls for whether the Proposer or Responder previously participated in game experiments. The ultimatum estimator estimates the willingness to pay of the Proposer and Responder simultaneously. The upper half of the table shows estimated effects of the independent variables and covariates on the Proposer's willingness to pay, and the lower half estimates of effects on the Responder's willingness to pay. For example, in model 1, the Proposer has a lower willingness to pay when the responder has a high capacity, meaning the Proposer would like the Responder to pay more of the costs of GHG mitigation. All models were fit with an intercept

P Proposer, *R* Responder, *HR* historical responsibility

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

between the Responder's capacity/historical responsibility and her level of altruism, risk aversion, or ecological concern. Findings remained unchanged with the inclusion of demographic variables (see SI-6).

Contrary to our expectations, it does not matter whether historical responsibility and capacity is endogenously (i.e., acquired through players' own actions) or exogenously determined (i.e., randomly assigned). Thus, even when individuals "inherit" capacity and historical responsibility, they still feel responsible. We find that the more willing an individual is to take risks, the lower her willingness to pay. As expected, the more altruistic an individual is, the higher her willingness to pay for the costs of climate change mitigation. Counter to our expectations, a participants' ecological concern does not have a significant effect.

5 Discussion and conclusions

As noted in the introduction, we are interested in whether and to what extent the public cares about the principles of historical responsibility and capacity when forming preferences concerning climate change mitigation and in whether governments should continue to use these as justifications for more ambitious climate change mitigation measures.

Our results show that, unless a Proposer is highly altruistic, her capacity and historical responsibility do not affect her willingness to pay for climate change mitigation. The Responder's capacity and historical responsibility, in contrast, significantly affect the Proposer's willingness to pay regardless of her level of altruism. Thus, as Proposers, individuals seem to selectively apply fairness principles to the other player and not to themselves. This holds even when controlling for participant characteristics such as age, education, income, political ideology, and start time. One explanation for this phenomenon could be a self-serving bias. Such bias occurs when people misevaluate and justify their own or other behaviors in a self-serving manner. This means that there is a tendency to take personal responsibility for desirable outcomes and externalize responsibility for undesirable ones (Campbell and Sedikides 1999). We find some evidence for this attribution bias in participants' post-round explanations of their decisions (see SI-7 for all comments). For example, in one round of the game, a Proposer offered to pay 7 ECU and the Responder rejected. The Proposer argued, *because my contribution of risk was about 7%*, while the Responder wrote, *I shouldn't have to pay more when he has more ECU*. These statements suggest that the Proposer focused on the contribution to climate risk, which cast him in a more favorable light, instead of capacity. Further research should investigate whether such self-serving bias does, in fact, explain the Proposer's behavior observed in our experiment.

As Responders, participants' willingness to pay is significantly affected by their capacity and historical responsibility but not the Proposer's, which holds even when controlling for participant characteristics. This behavioral pattern could be due to the design of the ultimatum game. In Responders' post-round explanations, participants frequently stated that they accepted an offer in order to avoid a climate catastrophe. Thus, if Proposers base their decisions on the attributes of the Responder, and Responders tend to accept offers to avoid a catastrophe, it could appear as though Responder's are only affected by their historical responsibility and capacity. This is unlikely since the results remain unchanged when analyzing only the rounds where the Responder rejected the Proposer's offer. However, to test for this, further research could modify the experimental design and add the possibility of a climate catastrophe even if the costs of mitigation are paid.

Another explanation for the Responder's behavior is the structure of ultimatum game. One could criticize that one player proposing and another responding is artificial. However, we believe that the ultimatum game is an appropriate reflection of reality. Even if there are multiple countries involved in the negotiations, not all countries possess the same weight to other countries. The USA has by far the largest weight for countries like India and China, and in a sense, what the USA does is an ultimatum to China and India. The USA and its citizens, in turn, respond to the commitments of China and India. Even though the actual climate negotiations are not formalized as an ultimatum game, in practice, climate negotiations turn into one.

In either position, however, participants' choices as well as their post-round comments clearly indicate that concerns about fairness play a key role when they form preferences concerning climate change mitigation and their willingness to pay for it. For example, one

Proposer stated, *I thought it was fair of me*, while the Responder claimed, *It was fair, eliminated chance for catastrophe and ended up with the same endowment as the other player.*

Our results are somewhat at odds with recent work that suggests the mass public is willing to support unilateral climate policies, meaning undertaking mitigation regardless of whether other countries undertake similar efforts (Tingley and Tomz 2013, Bernauer and Gampfer 2015). The main explanation for these seemingly conflicting preferences of individuals is likely to pertain to the context. We suggest that the context of international bargaining tends to bring out individuals' desire for reciprocity and fairness. Our experiment is in fact a stylized form of international negotiation over climate policy where the distribution of costs of GHG mitigation is explicit and thus clearly at the forefront of bargaining, whereas the distribution of mitigation costs between countries remains rather vague in survey experiments on unilateralism (see Bernauer and Gampfer 2015, Tingley and Tomz 2013).

If it is the context of explicit bargaining over mitigation costs that brings out individuals' preferences for fairness, public support for climate policy is likely to be stronger if policymakers frame their INDCs more in terms of unilateral policies serving their respective country's interests rather than as a contribution to a global burden sharing effort. This interpretation lines up well with recent research showing that climate change narratives around national identity can contribute to enhancing public support for mitigation efforts.[10] Further research could examine this argument more systematically by combining framing experiments with interactive games in order to find out to what extent public support for climate policy differs across the two contexts.

A note of caution, although policymakers have considerable room to maneuver in framing their respective national mitigation efforts (INDCs) as an independent unilateral effort, there are obvious limits to this (see Bernauer and McGrath 2016) since geophysical consequences of the problem are materializing on a global scale. This implies that states and citizens will continue to resort to fairness principles to some extent. This also means that the restructuring climate governance from centralized target setting to the present decentralized nonbinding process has not resolved the issue of fairly distributing mitigation costs.

Our results suggest a different restructuring of the climate governance framework might resolve these issues. If the UNFCCC publicized *all* intended commitments at the same time, each country would initially be in the position of Proposer. After the UNFCCC simultaneously publishes all countries commitments, there could be an adjustment period allowing countries to alter their commitments. This means each country is essentially a Responder. After which, the UNFCCC could release the revised commitments prior to the climate negotiations. This might bolster countries to make more ambitious commitments and recognize their capacity and historical responsibilities.

In summary, we should expect the global burden sharing debate to resurface even within the bottom-up climate governance framework and become more contentious as commitments become more ambitious. However, our results also suggest that the new INDC structure can be useful in freeing up some political space for more ambitious national mitigation efforts, in the sense of allowing policymakers to leverage public support for unilateral measures and overcome parochial domestic interests that seek to block national efforts with reference to insufficient efforts by other countries.

[10] Climate Outreach, http://climateoutreach.org/ (Accessed 1 December 2016)

Acknowledgements The research for this article was funded by the ERC Advanced Grant 'Sources of Legitimacy in Global Environmental Governance' (Grant 295456) and supported by ETH Zürich. We are grateful to Robert Gampfer, Mike Hudecheck, Vally Koubi, Liam McGrath, Lionel Miserez, Irina Shaymerdenova, and Florian Schmidt.

References

Almås I, Cappelen AW, Sørensen EØ, Tungodden B (2010) Fairness and the development of inequality acceptance. Science 328:1176–1178

Andreoni J, Miller J (2002) Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. Econometrica 70(2):737–753

Balietti, S (2016). nodeGame: real-time, synchronous, online experiments in the browser. Behav Res Methods: 1–20

Barabas J, Jerit J (2010) Are survey experiments externally valid? Am Polit Sci Rev 104:226–242. doi:10.1017/S0003055410000092

Barrett S (2003) Environment and statecraft: the strategy of environmental treaty-making. Oxford University Press

Barrett S (2006) Climate treaties and "breakthrough" technologies. Am Econ Rev 96:22–25

Barrett S (2011) Avoiding disastrous climate change is possible but not inevitable. P Natl Acad Sci USA 108: 11733–11734

Bechtel MM, Scheve KF (2013) Mass support for global climate agreements depends on institutional design. P Natl Acad Sci USA 110:13763–13768

Bernauer T, Gampfer R (2015) How robust is public support for unilateral climate policy? Environ Sci Pol 54: 316–330

Bernauer T, Gampfer R, Kachi A (2014) European unilateralism and involuntary burden-sharing in global climate politics: a public opinion perspective from the other side. Eur Union Polit 15:132–151. doi:10.1177/1465116513496878

Bernauer T, McGrath LF (2016) Simple reframing to boost public support for climate policy. Nat Clim Chang. doi:10.1038/nclimate2948

Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. Am Econ Rev 90:166–193

Buchan NR, Grimalda G, Wilson R, Brewer M, Fatas E, Foddy M (2009) Globalization and human cooperation. Proc Natl Acad Sci 106(11):4138–4142

Burton-Chellew MN, May RM, West SA (2013) Combined inequality in wealth and risk leads to disaster in the climate change game. Clim Chang 120:815–830. doi:10.1007/s10584-013-0856-7

Campbell WK, Sedikides C (1999) Self-threat magnifies the self-serving bias: a meta-analytic integration. Rev Gen Psychol 3:23–43

Cappelen AW, Nielsen UH, Sørensen EØ, Tungodden B, Tyran JR (2013) Give and take in dictator games. Econ Lett 118:280–283

Carlsson F, Kataria M, Lampi E, Löfgren Å, Sterner T (2011) Is fairness blind?—the effect of framing on preferences for effort-sharing rules. Ecol Econ 70:1529–1535

Carlsson F, Kataria M, Krupnick A, Lampi E, Löfgren Å, Qin P, Sterner T (2013) A fair share: burden-sharing preferences in the United States and China. Resour Energy Econ 35:1–7

Dunlap RE, Van Liere KD, Mertig AG, Jones RE (2000) New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale. J Soc Issues 56:425–442

Dutta PK, Radner R (2009) A strategic analysis of global warming: theory and some numbers. J Econ Behav Organ 71:187–209

Fehr E, Schmidt KM (1999) A Theory of Fairness, Competition, and Cooperation. The Quarterly J of Econ 114(3):817–868

Fuglestvedt JS, Kallbekken S (2016) Climate responsibility: fair shares? Nat Clim Chang 6:19–20

Gampfer R (2014) Do individuals care about fairness in burden sharing for climate change mitigation? Evidence from a lab experiment. Clim Chang 124:65–74. doi:10.1007/s10584-014-1091-6

Gintis H, Bowles S, Boyd R, Fehr E (2003) Explaining altruistic behavior in humans. Evol Hum Behav 24(3): 153–172

Huff C, Tingley D (2015) Who are these people? Evaluating the demographic characteristics and political preferences of MTurk survey respondents. Res & Polit 2:1–12

Paolacci G, Chandler J, Ipeirotis PG (2010) Running experiments on Amazon Mechanical Turk. Judgm Decis Mak 5:411–419
Rogelj J, den Elzen M, Höhne N, Fransen T, Fekete H, Winkler H, Schaeffer R, Sha F, Riahi K, Meinshausen M (2016) Paris Agreement climate proposals need a boost to keep warming well below 2 °C. Nature 534:631–639
Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? J Econ Perspec 21:153–174
Mattoo A, Subramanian A (2012) Equity in climate change: an analytical review. World Dev 40:1083–1097
Milinski M, Röhl T, Marotzke J (2011) Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. Clim Chang 109:807–814
Milinski M, Sommerfeld RD, Krambeck HJ, Reed FA, Marotzke J (2008) The collective-risk social dilemma and the prevention of simulated dangerous climate change. P Natl Acad Sci USA 105:2291–2294
Murphy RO, Ackermann KA (2013) Social value orientation theoretical and measurement issues in the study of social preferences. Personal Soc Psychol Rev 18:13–41
Ramsay, K. W., & Signorino, C. S. (2009). A statistical model of the ultimatum game. University of Rochester
Schleich J, Dütschke E, Schwirplies C, Ziegler A (2016) Citizens' perceptions of justice in international climate policy: an empirical analysis. Clim Policy 16:50–67
Sears DO (1986) College sophomores in the laboratory: influences of a narrow data base on social psychology's view of human nature. J Pers Soc Psychol 51:515–530
Signorino CS, Yilmaz K (2003) Strategic misspecification in regression models. Am J Polit Sci 47:551–566
Tavoni A, Dannenberg A, Kallis G, Lšschel A (2011) Inequality, communication, and the avoidance of disastrous climate change in a public goods game. P Natl Acad Sci USA 108:11825–11829. doi:10.1073/pnas.1102493108
Tingley D, Tomz M (2013) Conditional cooperation and climate change. Comp Polit Stud 20:1–25. doi:10.1177/0010414013509571
Underdal A, Wei T (2015) Distributive fairness: a mutual recognition approach. Environ Sci Pol 51:35–44
Wintre GM, North C, Sugar LA (2001) Psychologists' response to criticisms about research based on undergraduate participants: a developmental perspective. Can Psychol 42:216
Zizzo DJ (2010) Experimenter demand effects in economic experiments. Exp Econ 13:75–98