How Social Preferences Influence the Stability of a Climate Coalition

Lin, Yu-Hsuan

Department of Economics, the Catholic University of Korea

March 2018

How Social Preferences Influence the Stability of a Climate Coalition

Yu-Hsuan Lin[1]

Department of Economics, the Catholic University of Korea, South Korea

Version: July 2018

[Abstract]

This study examines the impact of social preferences on the individual incentives of participating in climate coalitions with laboratory experimental evidence. The theoretical result suggests that, when players are self-interested, dominant strategy equilibrium could exist conditionally. Players could be either critical or non-critical to an effective coalition. Their dominant strategy, either joining or not joining, depends on its contribution to total abatement. However, inequality-averse individuals may reshape the coalition formation. The laboratory evidence in this study supports that most players were inequality-averse and the coalition size was usually larger than the dominant strategy equilibrium and unstable. Nevertheless, the inequality-averse attitude was positively associated with the incentives of participation. Particularly, when they were non-critical players, egalitarians were likely to give up the free riding benefit by joining a coalition. Our findings help to understand the climate coalition formation.

[Key words]: experimental design; social preference; inequality-aversion; international environmental agreements; climate coalition

1. Introduction

Since Barrett (1994), a large number of studies (such as Bahn, Breton, Sbragia, and Zaccour (2009), Barrett (2001), Bratberg, Tjøtta, and Øines (2005), Eyckmans and Finus

[1] 43 Jibong-ro, Wonmi-gu, Bucheon-si, Gyeonggi-do, 14662, Republic of Korea.
E-mail: yuhsuan.lin@catholic.ac.kr

This work was supported by the Catholic University of Korea Research Fund.

(2006) and Rubio and Ulph (2006)) have explored the formation of international environmental agreements (IEAs) on climate change. Most theoretical simulation on IEAs predicted, without policy instruments, the number of signatories of a stable IEA may be very small (Diamantoudi & Sartzetakis, 2006; Grüning & Peters, 2010). This is in sharp contrast to empirical evidence. Recent experimental studies (Burger & Kolstad, 2010; Kosfeld, Okada, & Riedl, 2009; McEvoy, Cherry, & Stranlund, 2014) have pointed out that actual coalition formations are usually larger than theoretical predictions. A growing number of studies have proposed that this challenge was due to the fundamental assumption of self-interest (Hoel & Schneider, 1997; Willinger & Ziegelmeyer, 2001).

The self-interest has been widely employed in the majority of studies of IEAs. There appear to be incentives for countries to pursue their self-interest (Breton, Sbragia, & Zaccour, 2010). Several studies (Charness & Rabin, 2002; Dannenberg, Löschel, Paolacci, Reif, & Tavoni, 2015; Grüning & Peters, 2010; Hadjiyiannis, İriş, & Tabakis, 2012; Lange, 2006) had suggested that this assumption was not enough to explain actual individual decision makers' behaviours in public goods games. The role of social preferences (also known as other-regarding preferences) has been proposed to address this limitation. Several survey studies had also identified individuals' social values and norms on preferences may influence climate policies (Alló & Loureiro, 2014; Domínguez Arcos, Labandeira Villot, & Loureiro García, 2011; Hanemann, Labandeira, & Loureiro, 2011; Svenningsen, 2017; Svenningsen & Thorsen, 2017). Also, İriş, Lee, and Tavoni (2016) and Lin (2017) support with experimental evidence that the contributions in the climate coalition are subject to public pressure, instead of the self-interest. To reach higher environmental standard for climate change mitigation, people with higher social values were willing to contribute more.

As experimental economics provided some evidence for more complex human behaviour, extending the theory of IEAs to a broader class of preferences is clearly promising. Most of them, such as Kosfeld et al. (2009) and Grüning and Peters (2010), suggested that countries' preferences incorporate justice and fairness could promote abatement and enlarge the coalition size. On the other hand, Kolstad (2014) argued that social preference may reduce the equilibrium size of a coalition. Although the influence of social preferences on coalition formation has be overviewed, the influence on *individual incentives* for participating in a coalition has not yet been properly explored. These studies used the design of multiple equilibria, where agents have no dominant strategy, would limit the prediction of individuals' decisions. Hence, this study attempts to answer two questions: Does the concern about payoffs of others affect to individuals' decisions? How does the individual social preference influence the coalition formation?

This study investigates the influence of individual social preferences in a climate coalition with laboratory experimental evidences. A theoretical model was built to illustrate individual preferences in a dominant strategy equilibrium. When players are inequality-neutral, based on their weakly dominant strategies, their preferences could be either joining or not joining a coalition. In other words, they could be either critical or non-critical to an effective coalition. However, when players are inequality-averse, such clear-cut preferences and a stable coalition may not exist. There would be three possible outcomes: first, if players are weakly inequality-averse, the outcome remains the same as the inequality-neutral prediction. Second, when critical players are strongly inequality-averse, these players would break an effective coalition. Third, when non-critical players are inequality-averse, these players would enlarge an effective coalition. This study verifies the theoretical predictions with a two-phase laboratory experiment: the first phase identifies individual inequality-averse attitudes and the second phase deliberately imposes treatments on a group of subjects in the interest of observing their preferences.

There are two primary strengths in this study. First, individual preferences could be clear-cut and predictable in the model. This experiment design would provide a suitable environment to observe individual preferences when every player has a weakly dominant strategy. Second, in order to verify the theoretical prediction, a laboratory experiment was conducted. It provides detailed observations on the process of individual decision and coalition formation.

The outline of the study is as follows: section 2 builds a benchmark model and an inequality-averse model. Based on the theory, section 3 introduces two experiments. Then, Section 4 reports the experimental results and implications. The final section concludes.

2. The Models

Consider a simple climate coalition game, N heterogeneous countries consider their participation in a climate coalition. We assume that the marginal benefit of each country, γ , is in the range of 0 to 1 and the abatement cost is standardized in the range of 0 and 1. Suppose that n countries join while the others do not participate in a coalition. Signatories would do fully abate whilst nonsignatories would do fully pollute. Two scenarios are considered: the first scenario assumes that countries are self-interested and the second scenario assumes that countries have social preferences.

Let $u_i\ (n)$ denotes the welfare of country i if it is a member of a n-member coalition and $u_j\ (n)$ denotes the welfare of country j if it is outside a n-member coalition. Following d'Aspremont, Jacquemin, Gabszewicz, and Weymark (1983), a self-enforcing coalition of n^* satisfies the following constraints:

$$u_i\ (n^*) > u_j\ (n^* - 1) \quad \text{(Eq.1)}$$

$$u_j\ (n^*) > u_i\ (n^* + 1) \quad \text{(Eq.2)}$$

The internal constraint (Eq.1) denotes that a signatory has no incentive to leave the n^* member coalition and becomes a non-signatory. The external constraint (Eq.2) indicates that a non-signatory have no incentives to participate in a coalition as the $(n^* + 1)$-th member. When both constraints are satisfied, a n^* member stable coalition would exist.

2.1 Benchmark model with self-interest preference

When countries are self-interested, following Lin (2017), a country's welfare function is equal to the country's own payoff which depends on the number of signatories and its membership status. In a profitable n-member coalition, non-signatories do full pollute. The payoff of a non-signatory j (π_j) is the product of its marginal benefit rate and the total abatement from signatories. On the other hand, all signatories work as one to do full abatement and pay the standard cost. The coalition payoff (Π_s) is the joint payoffs of n members $(\pi_i, \forall i = 1, \dots, n)$ and can be written as $\Pi_s = \sum_{i=1}^{n} \pi_i = \sum_{i=1}^{n} [(-1) + \gamma_i n]$. Signatories share the coalition payoff equally among signatories. Hence, the payoffs of a nonsignatory j and a signatory i are written respectively as

$$\pi_j = \gamma_j n \quad \text{(Eq.3)}$$

$$\pi_i = (-1) + \sum_{i=1}^{n} \gamma_i \quad \text{(Eq.4)}$$

Taking (Eq.3) and (Eq.4) into the internal and external constraints (Eq.1) and (Eq.2), a stable coalition exist as the constraints are satisfied. Though the formation of coalition is predictable, multiple stable coalition combinations existed in the past experimental studies, such as Kosfeld et al. (2009) and Burger and Kolstad (2010). In other words, the size of coalition might be predictable, individual decisions were not. Due to no clear-cut preference for countries, these studies failed to foresee individual decisions in the membership game. This study, in order to have better prediction on individual decisions, focus on the cases of dominant strategy equilibrium. To reach such unique equilibrium, an additional condition should be satisfied:

$$\gamma_{n^*} > \sum_{j=n^*+1}^{N} \gamma_j \quad \text{(Eq.5)}$$

These constraints and condition categorise countries into two groups: *critical* and *non-critical* countries. Critical countries, with larger marginal benefits, are essential to an effective coalition. Therefore, the weakly dominant strategy of critical players is to participate in a coalition. Non-critical countries, with smaller marginal benefits, can contribute the coalition but not necessary. The weakly dominant strategy of non-critical players is not to participate.

The condition (Eq.5) also implied that any critical country would not be replaced by all of the non-critical countries. The condition ensures that the coalition is the only stable combination to be profitable. In other words, critical countries would participate in a coalition because they were necessary members and non-critical countries would not participate because they could take advantage from free-riding. While we acknowledge this is indeed a strong condition, in order to identify the individual incentives to participate in the coalition, this condition provided better observation of the individual decisions in the membership game.

2.1 Inequality-averse preference in a coalition game

Let us now consider the scenario of inequality-averse preferences. Different from (Eq. 3) and (Eq.4), the welfare of a country k ($k \in [1, N]$) depends on not only its own payoff but also the payoff gaps between its and others. Following Fehr and Schmidt (1999), the magnitude of inequality-aversion indicates the level of dislike for unfair outcomes. The inequality-averse welfare of a country k

$$u_k(n) = \pi_k - \frac{\alpha_k}{N-1}\sum_{k'} \max\left(\pi_{k'} - \pi_k, 0\right) - \frac{\beta_k}{N-1}\sum_{k'} \max\left(\pi_k - \pi_{k'}, 0\right) \quad \text{(Eq.6)}$$

where k' is any other country except k. The first term is the payoff of country k. The second and third terms represent the average payoff gap from the other country k' with the disadvantage-loss parameter $\alpha_k$and advantage-loss parameter β_k, respectively. Both of the parameters presented an inequality-averse magnitude of k and were between 0 (inequality-neutral) and 1 (strongly inequality-averse). Due to the abatement cost, a critical country earns a lower payoff than what a non-critical player has. A critical country faces the disadvantage-loss only. On the other hand, depends on its marginal benefit, a non-critical country may have both advantage-loss and disadvantage loss.

When inequality-aversion is taken into the internal and external constraints (Eq.1) and (Eq.2), the coalition formation depends on the individual inequality-averse magnitudes. The coalition formation could become either a stable n^*-member coalition, unstable, or a stable coalition larger than n^*. When all of the countries were inequality-neutral or weakly inequality-averse, a stable n^*-member coalition exists as the self-interested outcome. When a critical country was strongly inequality-averse, the country feels disadvantaged from the payoff gap between players. The internal constraint has been changed by the absence of such country from the effective coalition. Nevertheless, the country has the incentive to participate if everyone yields nothing from a collapsed coalition. Therefore, the coalition formation became unstable. In the last circumstance, when any non-critical country was strongly inequality-averse, that country could participate to mitigate the payoff gap. Therefore, the external constraint has been changed so that the coalition size is expanded and larger than n^*.

Intuitively, there were a number of effects with inequality-aversion. First, egalitarianism reduces the individual welfare when the payoffs among the countries were not equal. A coalition could be enlarged by a non-critical egalitarian country, when it sought for smaller advantage loss. Second, the transfer mechanism where signatories equally shared the coalition payoff could minimise the payoff gap among the countries. However, except for a grand coalition, signatories always suffered the disadvantage loss from non-signatories. An expanding IEA tended to exacerbate the payoff gap between the signatories and the non-signatories. Egalitarian signatories could punish free-riders behaviour by turning down an effective IEA. In other words, the effects of inequality-aversion could shape the stability and the formation of IEAs both internally and externally.

3. Experiment Design and Procedure

The experiment was conducted at the Centre for Experimental Economics (EXEC) laboratory at the University of York (UK) and programmed with z-Tree (Fischbacher, 2007). Fifty subjects were invited through the Online Recruitment System (Greiner, 2004). They were students from different countries and studied various disciplines. In order to understand the coalition formation, we mimicked the diversity in the real world where decision makers have different nationalities and multidisciplinary knowledge in this experiment.

A pre-experimental questionnaire gathered demographic information, including the subjects' degree disciplines, age (the year they were born), ethnicity, political orientation, and their level of belief in a religion. Another two questions collected information about their self-evaluated preferences. The question regarding religion identified the subjects' belief attitude on a scale ranging from 1 (not religious at all) to 5 (extremely religious). The distribution of the level of religious attitude showed that most subjects considered themselves as mild belief. The last question aimed to indicate the subjects' political preference (level one indicates left, level two indicates centre-left, level three indicates neutral, level four indicates centre-right, and level 5 indicates right). The distribution showed that most respondents were pro-left wingers.

In order to ensure data quality, the subjects had to comprehend the rules of the game as much as possible. They were not allowed to exchange information and no conversation was allowed (except for asking the experimenter to clarify the questions) during the experiment. The experimenter introduced the rules and gave the participants time to read through the instructions thoroughly and to accomplish the controlled questions. At the beginning of each part of the experiment, four control questions were asked in order to test the subjects' understanding. The experiment started when all of

the subjects had answered all of the control questions correctly. It was comprised of two phases with the following design.

3.1 An inequality-averse preference test

In this test, we aimed to examine the individuals' attitudes towards inequality-aversion. In order to extract information from a purified environment, the subjects were paired without knowing their partners or their partners' decisions. Each subject had two roles: dictator and receiver. A receiver passively earned allowance from the dictator's decision. A dictator, on the other hand, decided to share a £5 allowance with his/her receiver. There were two ways to share as shown in Table.1. Option 1 shared the allowance equally, while option 2 allocated the allowance unjustly with an all-or-nothing allocation at a certain probability.

Table.1. Inequality-aversion Test

Round	Option 1	Option 2
1	(£2.5, £2.5) for sure	(£0, £5) with probability 0%; (£5, £0) with probability 100%
2	(£2.5, £2.5) for sure	(£0, £5) with probability 10%; (£5, £0) with probability 90%
3	(£2.5, £2.5) for sure	(£0, £5) with probability 20%; (£5, £0) with probability 80%
4	(£2.5, £2.5) for sure	(£0, £5) with probability 30%; (£5, £0) with probability 70%
5	(£2.5, £2.5) for sure	(£0, £5) with probability 40%; (£5, £0) with probability 60%
6	(£2.5, £2.5) for sure	(£0, £5) with probability 50%; (£5, £0) with probability 50%
7	(£2.5, £2.5) for sure	(£0, £5) with probability 60%; (£5, £0) with probability 40%
8	(£2.5, £2.5) for sure	(£0, £5) with probability 70%; (£5, £0) with probability 30%
9	(£2.5, £2.5) for sure	(£0, £5) with probability 80%; (£5, £0) with probability 20%
10	(£2.5, £2.5) for sure	(£0, £5) with probability 90%; (£5, £0) with probability 10%
11	(£2.5, £2.5) for sure	(£0, £5) with probability 100%; (£5, £0) with probability 0%

Option 1 was a fair allocation where the dictator faces no unfair loss. On the other hand, the all-or-nothing allocation in option 2 indicated two extreme cases. As described in (Eq.6), an inequality-averse agent considered both advantage-loss and disadvantage-loss. The range of the two extremities could be normalised. The range between two extreme unfair outcomes was normalised, so both the advantage- and disadvantage-losses could be merged as one inequality-averse indicator. Although a subject might suffer more from disadvantage than advantage, two reasons supported this technique. In practice, it is not easy to find a subject's preference without standardising the unit of the utility. In the literature, the experimental evidence showed that the disadvantage factor was not necessarily smaller than the advantage factor (Dannenberg, Riechmann, Sturm, & Vogt, 2012; Yang, Onderstal, & Schram, 2016). In round 1, the all-or-nothing

allocation would be taken by a rational subject because the outcome was definitely better than that from option 1. By contrast, in the final round, the outcome for the fair allocation was better than that for the all-or-nothing allocation. For each subject with a consistent preference, there existed a point with a certain probability where the subject would switch from the all-or-nothing allocation to the fair allocation. The switch point indicated the individuals' attitudes toward inequality-aversion.

If the subjects were self-interested, then their welfares were the same to their monetary payoffs. In other words, they would switch when the expected outcome of all-or-nothing allocation was equal to that of fair allocation. If subjects were inequality-averse, their utilities were lower than their monetary payoffs. They were more likely to take an equal allocation in order to avoid extremely unfair consequences. Inequality acceptors, which never chose an equal allocation, could be possible, but they were uncommon in reality (as seen in the experimental results later). They could be captured in this experimental design. Therefore, we excluded those inequality acceptors from our analyses, similar to (Fehr & Schmidt, 1999).

It is important to bear in mind that this test could be characterised by strategic uncertainty due to the fact that a series of probabilities were involved. The subjects' risk attitudes might have been involved in their decisions. In other words, it might have been difficult to distinguish the risk aversion and inequality aversion in this study. This issue might be avoided by employing two separate games in order to indicate the attitudes toward disadvantage- and advantage-aversion, such as those developed by Blanco, Engelmann, and Normann (2011) and (Yang et al., 2016). However, this study was superior for two reasons. First, the two-games created another bigger issue in that the measurement of two attitudes might have been biased. Second, there was a significant positive correlation between the inequality-aversion and risk-aversion (Carlsson, Daruvala, & Johansson-Stenman, 2005; Kroll & Davidovitz, 2003). It was unnecessary to distinguish the inequality-aversion from the risk-aversion.

3.2 Coalition game

This public goods game mimicked the climate coalition conventions. The subjects were assigned different roles in a group of five anonymous persons for the entire session. As described in Equations (1) and (2), the payoffs depended on the marginal benefit of the total abatement. In this study, we built eight treatments of various marginal benefits. Each group played four treatments. Each treatment had 2 to 4 critical players whilst the rest played a role of non-critical. As explained earlier, based on the assumption of self-interest, the unique-equilibrium design could help to identify individual decisions. As illustrated earlier, critical players were essential for a profitable coalition, while non-critical players had the incentive to free ride.

When the subjects had strong inequality-averse attitudes, then the critical players might have had the incentive to break the coalition internally. On the other hand, non-critical players might have given up the free-riding benefit by participating in a coalition. In this study, we assigned each subject a particular payoff table, which contained all of the possible payoffs with the corresponding coalition combinations. The payoff depended on the given parameters and the coalition formation. For any unprofitable coalition, all of the subjects in the group gained nothing in return. The possible payoffs for the subjects ranged from £0 up to £24.

4. Experimental Results and Analyses

The results for the inequality-averse test demonstrated that 31 out of the 50 subjects had clearly switched from the all-or-nothing allocation to the fair allocation. In particular, 2 subjects stuck with the fair allocation for the entire session. Their behaviours indicated their individual attitude toward inequality-aversion.

Table.2. Number of Fair Allocation taken

Round	1	2	3	4	5	6	7	8	9	10	11
Number for Fair Allocation Taken	10	11	8	23	33	35	48	48	47	48	46

Table 2 presents the number of fair allocations taken in each round. Initially, most of the subjects preferred the all-or-nothing allocation. Then, their decisions switched to the fair allocation. When the expected payoff of all-or-nothing became lower than that of the fair allocation, it is unsurprising that almost everyone took the fair allocation.

Regarding the coalition formation in the membership game, effective coalitions were formed in 387 out of 600 rounds, and the formation was usually larger than the self-interested equilibrium size. The actual coalition formation matched the self-interested equilibrium in only 112 rounds. The coalitions were usually neither stable nor convergent to a particular coalition. With the same treatments, the coalition formation varied in different groups. For example, group 6 and group 8 both took treatments 5 to 8. Group 6 formed profitable coalitions in 47 rounds, but group 8 achieved profitable coalitions in only 12 rounds.

Table.3. Probit Estimations of Probability of Joining a Coalition

Variable	Probit MLE(1)	Probit MLE(3)	Probit MLE(5)
Constant term	15.00 (12.37)	4.81 (16.74)	21.56 (16.81)
Inequality-Averse Attitude	0.69 ** (0.17)	0.70 ** (0.26)	0.53 * (0.23)
Age	-0.008 (0.006)	-0.002 (0.008)	-0.01 (0.008)
Politic Attitude	0.09 ** (0.03)	-0.14 ** (0.05)	0.31 ** (0.04)
Religious Attitude	-0.06 ** (0.02)	0.05 (0.03)	-0.23 ** (0.03)
Critical player	1.16 ** (0.06)		
Marginal Benefit			-8.80 ** (1.07)
Past Total Contribution	0.57 ** (0.11)	0.97 ** (0.15)	0.03 (0.15)
Total Observations	2,520	1,400	1,120
Observations of Joining (% of total observation)	1,692 (67%)	1,185 (85%)	507 (45%)
Log Likelihood	-1339.29	-568.96	-697.09
Note: Each cell contains coefficient and standard error in parenthesis. *, ** are significant at 5% and 1% respectively.			

Turning now to the factors that might have affected the individual decisions, the maximum likelihood estimation of the binary probit regressions were employed as shown in Table 3. The variables included individual inequality-averse attitude (the times of taking the fair allocation), age (the year the subjects were born), the political attitudes from left to right, the religious attitudes from atheist to religious, the dummy variable of being critical players, the marginal benefit of the total contribution, and the group contribution in the previous round. Although the experimental design allowed for the existence of inequality lovers, which were players never chose an equal allocation. They were uncommon in reality as well as the experimental studies. Therefore, similar to Fehr and Schmidt (1999), five inequality lovers were excluded and 2,700 observations were used for our analyses.

The estimation of Probit MLE(1) covered the observations of 2,520 individual decisions which exclude the decisions in the first round of the treatment. Among these

observations, nearly 70% of them decided to participate in a coalition. The inequality-averse, political and religious attitudes, the dummy variable of being critical players, and the total contribution in the previous round had significant positive effects on the decisions. As we predicted, subjects usually chose the dominant strategy. The strongly inequality-averse subjects were more likely to participate in a coalition. While the coalition size was larger in the past round, subjects were more likely to cooperate. The result also suggests that pro-right-wingers were more likely to cooperate and religionists were not.

Probit MLE(2) examines the observations of the critical players who were essential to form an effective coalition. 85% out of the 1,500 observations participated in a coalition. Though the theoretical prediction suggested that egalitarians might break a coalition internally. In contrast, the experimental result shows that the internal stability was ensured by the subjects with stronger inequality-averse attitudes. This interesting result might be illustrated as inequality-aversers were more likely to cooperate while more people cooperated. This was due to the cooperation could mitigate the payoff gaps between players.

Interestingly, pro-left-wingers were more likely to cooperate when they were critical to an effective coalition. That being said, subjects had stronger incentives to form a profitable coalition when they were egalitarians or pro-left-wingers. The coalition size in the past round was positively associated with the cooperation of the critical players. There are two contradictory effects of a larger coalition. The first effect is a larger payoff gap between signatories and free-riders. This would cause more disadvantage loss. The second effect is the gaps between less players. This would make less loss.

The estimation of Probit MLE(3) assesses the non-critical players those did not have to cooperate. The results showed that the free-riding incentives were rejected for nearly half of the 1,120 observations. Again, egalitarians were more likely to compromise and cooperate. Subjects with stronger attitudes towards inequality-aversion were more likely to participate in a coalition. Apart from the inequality-averse attitudes, the estimation examines the factors of free-riding. Non-critical players receive free-riding incentives from two factors: its marginal benefit and the coalition size. Our finding claims that high marginal benefit could lead to low cooperation. It was intuitive that non-critical players were less cooperative due to higher free-riding incentives. In contrast to the experimental study of Burger and Kolstad (2010), this study does not support their earlier finding that high marginal benefit would increase a coalition size. On the other hand, another free-riding incentives from the past coalition size was insignificant to the willingness of participation.

There are more important policy implications from our results. Individual political and religious attitudes had significant effects on the individual cooperation. In contrast to their preferences when they were critical, the pro-left-wingers were less cooperative when they were non-critical. On the other hand, the atheists were more likely to cooperate when they were non-critical.

5. Discussion and Conclusions

This is the first investigation that examines the impacts of social preferences on the individual incentives of participating in climate coalitions by using laboratory evidence. We consider a case of dominant strategy equilibrium which players could be either critical or non-critical to an effective coalition. A critical player was essential and cannot be replaced by all non-critical players. Theoretically, when countries with social preferences, egalitarians could reshape the coalition internally and externally. The coalition size might remain as the dominant strategy equilibrium, be enlarged, or unstable. Though this study was limited in the sample scale and specific case of dominant strategy solution, but it confirms causal links to individual participation in a climate coalition from inequality-averse, politic and religious attitudes, the free-riding incentive, and the past coalition contribution.

Turning back to answer the research questions, the coalition formation could be influenced by individual social preferences. This study suggests that the inequality-averse attitudes had significant positive association with the incentives of participation. In particular, when the subjects were non-critical players, the egalitarians were likely to give up the free-riding benefit by joining a coalition. This result could explain why the coalition formation was usually larger than the Nash equilibrium.

This study has multiple implications for public opinion elicitation and public policy. Our findings suggest that it is important to highlight not only the individual payoff but also the gap between players. Some significant factors could be illustrated intuitively. For instance, the past coalition size and the marginal benefit were associated with the incentives for critical and non-critical players. Apart from that, other factors may not have been intuitive. For example, the pro-left-wingers behaved strategically: they were cooperative when they were critical and less cooperative when they were non-critical to an effective coalition. In addition, those atheists were more cooperative when they were non-critical.

In conclusion, individual concerns for others do influence the coalition formation. No matter whether they were critical or not, inequality-aversion could lead to more participation in a coalition. The results of this study also suggested that the individual motivation could be affected by their political and religious attitudes. The implications could advise policy makers on constructing a solid climate coalition for a better future.

Ethical Compliance

1) Sources of Funding

Source of Funding: This study was funded by the Catholic University of Korea.

2) Potential Conflicts of Interest

Disclosure of Interest: The authors declare that they have no conflict of interest.

3) Informed Consent

Informed consent was obtained from each individual participant involved in this study.

4) Statement of Human Rights

This study was conducted in accordance with the 1964 Declaration of Helsinki and its subsequent amendments.

Reference

Alló, M., & Loureiro, M. L. (2014). The role of social norms on preferences towards climate change policies: A meta-analysis. *Energy Policy, 73*, 563-574.

Bahn, O., Breton, M., Sbragia, L., & Zaccour, G. (2009). Stability of international environmental agreements: an illustration with asymmetrical countries. *International Transactions in Operational Research, 16*(3), 307-324.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford economic papers*, 878-894.

Barrett, S. (2001). International cooperation for sale. *European Economic Review, 45*(10), 1835-1850.

Blanco, M., Engelmann, D., & Normann, H. T. (2011). A within-subject analysis of other-regarding preferences. *Games and Economic Behavior, 72*(2), 321-338.

Bratberg, E., Tjøtta, S., & Øines, T. (2005). Do voluntary international environmental agreements work? *Journal of Environmental Economics and Management, 50*(3), 583-597. doi:https://doi.org/10.1016/j.jeem.2005.03.002

Breton, M., Sbragia, L., & Zaccour, G. (2010). A dynamic model for international environmental agreements. *Environmental and Resource economics, 45*(1), 25-48.

Burger, N. E., & Kolstad, C. D. (2010). *International Environmental Agreements: Theory Meets Experimental Evidence*.

Carlsson, F., Daruvala, D., & Johansson-Stenman, O. (2005). Are People Inequality-Averse, or Just Risk-Averse? *Economica, 72*(287), 375-396.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics, 117*(3), 817-869.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J., & Weymark, J. A. (1983). On the stability of collusive price leadership. *Canadian Journal of economics*, 17-25.

Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2015). On the provision of public goods with probabilistic and ambiguous thresholds. *Environmental and Resource economics, 61*(3), 365-383.

Dannenberg, A., Riechmann, T., Sturm, B., & Vogt, C. (2012). Inequality aversion and the house money effect. *Experimental Economics, 15*(3), 460-484.

Diamantoudi, E., & Sartzetakis, E. S. (2006). Stable international environmental agreements: An analytical approach. *Journal of public economic theory, 8*(2), 247-263.

Domínguez Arcos, F., Labandeira Villot, X., & Loureiro García, M. (2011). Climate Change Policies And Social Preferences In Galicia And Spain. *Revista Galega de Economía, 20*(1).

Eyckmans, J., & Finus, M. (2006). New roads to international environmental agreements: the case of global warming. *Environmental Economics and Policy Studies, 7*(4), 391-414.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics, 114*(3), 817-868.

Fischbacher, U. (2007). z-Tree: Zutich Toolbox for Ready-made Economic Experiments. *Experimental Economics, 10*, 171-178.

Greiner, B. (2004). *The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics*. Retrieved from

Grüning, C., & Peters, W. (2010). Can justice and fairness enlarge international environmental agreements? *Games, 1*(2), 137-158.

Hadjiyiannis, C., İriş, D., & Tabakis, C. (2012). International environmental cooperation under fairness and reciprocity. *The BE Journal of Economic Analysis & Policy, 12*(1).

Hanemann, M., Labandeira, X., & Loureiro, M. L. (2011). Climate change, energy and social preferences on policies: exploratory evidence for Spain. *Climate Research, 48*(2/3), 343-348.

Hoel, M., & Schneider, K. (1997). Incentives to participate in an international environmental agreement. *Environmental and Resource economics, 9*(2), 153-170.

İriş, D., Lee, J., & Tavoni, A. (2016). Delegation and public pressure in a threshold public goods game: theory and experimental evidence.

Kolstad, C. D. (2014). *International environmental agreements among heterogeneous countries with social preferences*. National Bureau of Economic Research.

Kosfeld, M., Okada, A., & Riedl, A. (2009). Institution Formation in Public Goods Games. *The American Economic Review, 99*(4), 1335-1355.

Kroll, Y., & Davidovitz, L. (2003). Inequality aversion versus risk aversion. *Economica, 70*(277), 19-29.

Lange, A. (2006). The impact of equity-preferences on the stability of international environmental agreements. *Environmental and Resource economics, 34*(2), 247-267.

Lin, Y.-H. (2017). The effects of inequality aversion on the formation of climate coalition: Theory and experimental evidence *Economics of International Environmental Agreements* (pp. 73-88): Routledge.

McEvoy, D. M., Cherry, T. L., & Stranlund, J. K. (2014). International environmental agreements with endogenous minimum participation and the role of inequality. In J. H. Todd L. Cherry, David M. McEvoy (Ed.), *Toward a New Climate Agreement: Conflict, Resolution and Governance, Routledge* (pp. 93-105): Routledge.

Rubio, S. J., & Ulph, A. (2006). Self-enforcing international environmental agreements revisited. *Oxford economic papers, 58*(2), 233-263.

Svenningsen, L. S. (2017). *Distributive outcomes matter: Measuring social preferences for climate policy*. University of Copenhagen, Department of Food and Resource Economics.

Svenningsen, L. S., & Thorsen, B. J. (2017). *Preferences for distributional impacts of climate policy*. Retrieved from

Willinger, M., & Ziegelmeyer, A. (2001). Strength of the social dilemma in a public goods experiment: an exploration of the error hypothesis. *Experimental Economics, 4*(2), 131-144.

Yang, Y., Onderstal, S., & Schram, A. (2016). Inequity aversion revisited. *Journal of Economic Psychology, 54*, 1-16.