RUHR
ECONOMIC PAPERS

Rafat Beigpoor Shahrivar
Ilka Duesterhoeft
Marco Rogna
Carla J. Vogt

A Mechanism of Proportional Contributions for Public Good Games

RUB

#990

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #990

Responsible Editor: Manuel Frondel

ISSN 1864-4872 (online) – ISBN 978-3-96973-156-7

Ruhr Economic Papers #990

Rafat Beigpoor Shahrivar, Ilka Duesterhoeft, Marco Rogna, and Carla J. Vogt

A Mechanism of Proportional Contributions for Public Good Games

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at **http://dnb.dnb.de**

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/96973156
ISSN 1864-4872 (online)
ISBN 978-3-96973-156-7

Rafat Beigpoor Shahrivar, Ilka Duesterhoeft, Marco Rogna, and Carla J. Vogt[1]

A Mechanism of Proportional Contributions for Public Good Games

Abstract

Public good games in coalitional form, such as the ones depicting international environmental agreements for the reduction of a global pollutant, generally foresee scarce levels of cooperation. The incentive to free ride, that increases for higher levels of cooperation, prevents the formation of stable coalitions. The introduction of other-regarding preferences, in the form of Fehr and Schmidt utility functions, enlarges cooperation, but still at suboptimal levels. The present paper considers a further possibility, namely the introduction of a mechanism through which the contributions of players to the public good are proportional to the average contribution of the other players abiding to the mechanism: proportional contributions. The mechanism is therefore rooted into reciprocity. By applying it to a standard abatement game parameterized on the RICE model, we show that the mechanism is in fact able to increase cooperation both under standard and under F&S preferences. Stability of the grand coalition is never reached, but potential internally stable grand coalitions are achieved under F&S preferences. The attainment of higher cooperation comes at the expense of the level of global abatement that is lower when proportional contributions are in place.

JEL-Codes: C72, D63, H41, Q54

Keywords: Coalitional game; cooperation; F&S preferences; public good; reciprocity

January 2023

1 All RUB. – All correspondence to: Carla J. Vogt, Bochum University of Applied Sciences, Am Hochschulcampus 1, 44801 Bochum, Germany, e-mail: carla.vogt@hs-bochum.de

1 Introduction

Voluntary cooperation in the field of international climate policy is urgently needed, but hard to achieve. Since climate protection is a global public good, strong free riding incentives exist. While reducing greenhouse gas emissions comes at a measurable cost for a country, the beneficial effects of its mitigation efforts on the global climate are tiny, leaving it a rational choice for a single country not to abate at all or to choose inefficiently low levels of abatement. Countries are caught in a social dilemma.

In economics and game theory, coalition models have been employed to systematically study the incentives for voluntary cooperation in climate protection. The pessimistic results from the above sketched out simple reasoning have, more or less, been confirmed within this more complex framework. In a nutshell, coalition models have shown that cooperation is not to be expected by rational acting players in international climate policy whenever it would be needed most urgently, i.e. whenever the potential global welfare gains from voluntary cooperation are large. In these cases, only small coalitions of two or three players have been shown to be stable, not improving much compared to the non–cooperative Nash equilibrium.

The results mentioned above depend on some few assumptions typically made in coalitional models that partly have been conquered in the past two decades. On the one hand, assumptions about players' preferences have to be made. Economic reasoning builds upon two base pillars: Rationality and selfishness. The economic model of man assumes players to behave as trying to maximize their own utility. This assumption has been criticized from the point of view of behavioural economics. Economic experiments have shown over and over again that actual behaviour is not consistent with the assumption of pure selfishness. There is huge experimental evidence showing that fairness considerations may influence behaviour of participants besides classical efficiency concerns. Players reject positive amounts of money in the ultimatum bargaining game (Güth et al., 1982), they make positive donations in the dictator game (Kahneman et al., 1986) and they persistently and voluntarily cooperate in social dilemma games like, e.g., public good games or simple Prisoner's Dilemma games

(Isaac et al., 1994).

The challenge from experimental economics has been answered by theory. The seminal papers by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) introduce inequality aversion as part of the motivation of players in bargaining and dilemma games. Inequality aversion then has been introduced into models of coalition formation in international environmental policy. Lange and Vogt (2003), for example, show that even the grand coalition can be stabilized as a Nash equilibrium; their result, however, still rests on the assumption of identical players. In recent papers, Vogt (2016), Rogna and Vogt (2020) and Rogna and Vogt (2022), the cooperation enhancing effect of inequality aversion has been demonstrated even in a context of heterogeneous players, that seems to be more appropriate since countries significantly differ w.r.t. costs and benefits from climate mitigation to a huge extent. However, the mentioned papers also identify an appropriate transfer scheme as a prerequisite for the formation of effective and stable coalitions: Without a redistribution of wealth from rich to poor players, inequality aversion turns out not to be the great leap forward, already pointing to transfers as another important institutional feature and necessary ingredient for successful and broad cooperation.

Besides preferences and transfers as a device for redistributing income, a third institutional feature is important for determining the outcome in coalition formation games: Players have to decide about their voluntary contribution to the global public good climate protection. Hence, a contribution mechanism has to be specified for each coalition. Typically, for the members of a coalition, collective welfare maximization is assumed, from which the according optimal contributions for each member can be derived. For non–members, individual utility maximization is assumed. Clearly, this assumption refers to the second pillar of the economic model of man: rationality. Although from an academic point of view the assumption of maximizing behaviour seems to be straightforward, a look into real world policy making quickly reveals that quite different rules or principles for defining duties in international treaties may be adopted. In international trade policy, e.g., the WTO (and the GATT as its predecessor) rely upon the principle of reciprocity, as explained in more detail in section 2.

The present paper aims at introducing the principle of reciprocity in the field of international climate policy, employing it as a means to determine the countries' contributions to the global public good climate protection, thereby replacing the contribution mechanism based upon collective welfare maximization. In particular, the goal of the paper is to study which effects on the stability of coalitions and on their respective outcomes in terms of public good provision and associated welfare this may have. We study these effects by still relying on the assumption of inequality averse players, since there is strong evidence, particularly from experimental economics, that this motive significantly influences players' choice in bargaining and cooperation games.

Our main finding is that a reciprocity based contribution mechanism may indeed increase participation in stable coalitions. Particularly, for the case of inequality averse players, even the grand coalition may be potentially internally stable. The broader participation, however, comes at the cost of lower aggregate levels of public good provision and welfare.

The paper is organized as follows. Section 2 reviews the related literature, in section 3, our theoretical model is set up while in section 4, results from numerical simulations of this model are summarized. The last section, as always, concludes.

2 Literature Review

Coalition formation models date back to the seminal papers of Barrett (1994), Carraro and Siniscalco (1993) and Chander and Tulkens (1995), building on the foundations of coalition theory laid in cartel theory by d'Aspremont et al. (1983), which developed the core concept of internal and external stability then applied also to coalition formation in the field of international environmental agreements. These early papers rest on the assumption of purely selfish and rational players, thereby leading to rather pessimistic results about the prospects of self–enforcing agreements. The main finding is summarized by the so–called Barrett paradox: Stable coalitions cannot be achieved whenever they were mostly needed from a welfare point of view, i.e. when potential welfare gains by voluntary cooperation are high. The reason is quite clear: Whenever cooperation creates high welfare gains, the temptation to free–ride also increases. Several attempts of improving upon this pure outcome have been made, including e.g. introducing transfer payments within coalitions, but without substantial changes in the outcomes[1].

Matters may change if non–standard preferences are introduced. Inspired by findings in the field of experimental economics[2] and the theoretical developments triggered thereby, some papers started considering inequality aversion as part of players' motivation. Lange and Vogt (2003) introduce inequality aversion according to Bolton and Ockenfels (2000) into a coalition game finding that even the grand coalition can be a Nash equilibrium. This result, however, depends on the assumption of identical players, which seems hard to justify in the case of climate change since countries heavily differ with respect to costs of mitigation as well as benefits, respectively. Introducing heterogeneity, Lange (2006) finds that cooperation is much harder to obtain. van der Pol et al. (2012) consider altruistic preferences in a coalition game, leading to much more favourable outcomes. Vogt (2016) studies inequality aversion à la Fehr and Schmidt (1999). Using numerical simulations of a two stage coalition formation game, she finds that large and welfare improving coalitions can be stabilized, provided appropriate transfer schemes are in place. Building on these findings, Rogna and Vogt (2020) derive a utility maximizing transfer scheme which turns out to be highly redistributive, shifting income from rich to poor players. Applying this scheme in a coalition game leads to strong effects on the size of stable coalitions (nine out of twelve players) as well as on collective welfare (more than 95% of the efficient level). Rogna and Vogt (2022) apply Fehr–Schmidt utility functions to the Nordhaus RICE model. On the one hand, optimal temperature increase turns out to be significantly lower compared to the case of standard preferences. On the other hand, medium sized stable coalitions (six

[1]Several variants of these early models have been provided: McGinty (2007), Pavlova and De Zeeuw (2013) and Bakalova and Eyckmans (2019) investigate the case of heterogeneous countries; Weikard et al. (2015) include a minimum participation constraint; Diamantoudi and Sartzetakis (2015) and Breton and Garrab (2014) examine the consequences of (evolutionary) farsightedness.

[2]Experimental evidence for inequality aversion as a motive driving human behaviour is provided, e.g., in Güth et al. (1982), Bolton and Ockenfels (2005), Fehr and Schmidt (2006) and Fischbacher and Gachter (2010).

out of 12 players) turn out to be potentially internally stable (PIS)[3], leading to provision levels of the public good climate protection of about 50% of the efficient amount. Taken together, these papers clearly show that (i) inequality aversion per se does not necessarily lead to higher rates of cooperation but (ii) if accompanied by properly designed transfers, a change in behavioural assumptions is well suited to paint a much more optimistic view of prospects for cooperation in the field of climate policy. The grand coalition of all countries, however, has not been stabilized in case of heterogeneous players in all of the mentioned papers.

Besides preferences and transfers, two further set screws can influence the outcomes in coalition games. First, the equilibrium concept used. Here, the present paper sticks to the usual concept of external and internal stability, which is a straightforward extension of the notion of a Nash equilibrium to a coalition setting. Second, the contribution mechanisms may have an impact. All papers utilizing coalition games, at least to our knowledge, use the contribution mechanism based on collective welfare maximization. Principally, there is no reason why to restrict the choice of a contribution mechanism in this way: Players engaged in negotiations on issues of international policy may choose whatever they value suitable for their ends. A principle with great appeal would be the principle of reciprocity. Reciprocity might be one of the most important and powerful ideas both underlying normative as well as positive theories of human behaviour. On the one hand, reciprocity is a core concept inspiring huge parts of ethical reasoning as evidenced by various concepts of the so–called golden rule in philosophy as well as in all world religions. On the other hand, it forms the basis for important areas of international policy. The WTO (and the GATT as its predecessor) contain reciprocity as a constitutive principle of international trade policy. Important aspects in international trade resemble incentives in international climate policy. In fact, countries are caught in a Prisoner's dilemma game when deciding about their optimal tariffs (in general, optimal tariffs are not zero due to the well–known terms–of–trade effect). All countries would benefit if all countries would liberalize trade, i.e. all

[3]Loosely speaking, a coalition is PIS if it can be stabilized by some appropriate transfer scheme.

countries would lower their tariff rates. But unilateral action on this issue is not individually rational: If a single country lowers its tariffs alone, it only worsens its terms–of–trade while all other countries – not lowering their own tariffs – can gain. Hence, the only Nash–equilibrium in this game is not to lower the tariffs. The reciprocity rule underlying WTO (and GATT) can be interpreted as an institutional device to solve this dilemma situation by allowing member countries to carry out well–tailored sanctions on other parties. In fact, the literature explains the overwhelming success of WTO and GATT exactly by this institutional feature[4] (Bagwell and Staiger, 1999, 2004, 2016). Moreover, Freund (2003) provides empirical evidence for reciprocity in a sample of 91 free trade agreements that have been negotiated since 1980. Interestingly, she finds that reciprocity is present when countries are similar, i.e. in North–North and South–South treaties. Further empirical evidence for the positive role of reciprocity in international trade policy can be found in Rhodes (1989) and Lamp (2017).

Taken together, both its normative appeal and its practical use in important areas of international relations like trade policy make the reciprocity principle an ideal candidate for applying it also to the field of international environmental policy[5]. Besides its obvious appeal, in the field of climate mitigation severe difficulties arise in defining precisely what a reciprocal contribution to climate protection would be. These difficulties clearly arise from the simple fact that countries widely differ w.r.t. their responsibility for climate change on the one hand and w.r.t. to their vulnerability on the other hand. Therefore, quite distinct notions of reciprocity in the case of climate mitigation are imaginable. The following section lays out the specific notion of reciprocity employed in this paper. We develop a simple proportional scheme where a player's contribution is required to be proportional to the average of all contributions, thereby capturing a constituent of reciprocity: A player's contribution has to increase when all other players increase theirs, reflecting kindness as a

[4]Keohane (1982) already mentioned in an influential paper dated 1982 that reciprocity is a widely accepted behavioural norm not only underlying GATT but also the Basics Principle Agreement from 1972.

[5]Alternatively, reciprocity could also be introduced at the level of players' preferences. This task is investigated in a paper by Nyborg (2018). She finds that the grand coalition can be a Nash equilibrium, given that reciprocity is sufficiently widespread in the players' population.

reward to kind behaviour. By the same token, a decrease of others' contributions will be followed by a decrease of the contribution of the corresponding player.

3 The Model

As mentioned in the introduction, our objective is to analyse the performances of the proportional contributions mechanism both in case of standard and in case of F&S preferences. Performances are measured in terms of size of stable and potentially stable coalitions and in terms of provision of the environmental good. The standard case without proportional abatement is used as touchstone.

3.1 A Standard Coalitional Game of Abatement

In order to accomplish our task, we need a coalitional model of public good provision. We rely on the coalitional abatement model of Barrett (1994) for its popularity, being one of the first games of this type to be proposed, and for its simplicity. In this model, abatement of emissions, a global pollutant treated as a public bad, assumes the role of a public good. In order to abate emissions, countries have to forgo consumption, therefore diminishing own payoff. The cost of abatement is quadratic (convex), implying that the marginal cost of abatement is a monotonically increasing function of the same level of abatement. Emissions, being a public bad, decrease the payoff of a player and, consequently, abatement, by reducing them, increases it. Clearly, the private abatement of each single country benefits all countries being emissions a global pollutant. The benefits of abatement for each player are a linear–quadratic function of global abatement (the sum of abatement of all players), where the linear term is positive and the quadratic one negative. This implies declining marginal benefits of abatement, or, in other words, benefits are a concave function of global abatement. Mathematically, we have the following equation defining the payoff of a generic player i, belonging to the set of players N:

$$\pi_i = b_i Q - s_i \frac{Q^2}{2} - c_i \frac{q_i^2}{2}, \qquad \forall i \in N; \tag{1}$$

where q_i is the private abatement of player i, $Q = \sum_{i \in N} q_i$ is global abatement and b_i, s_i and c_i are the parameters of the model. The latter, c_i, defines how costly is the private abatement of player i, b_i how strong are the benefits of abatement and s_i the curvature of the concave benefit function for the same player.

The game is coalitional implying that it features the possibility of forming coalitions (C), meant to be international agreements between certain groups of countries to coordinate abatement. The standard assumptions for this class of games are that only one coalition at time can be formed[6] and that players belonging to a coalition will act for maximizing the welfare of the same coalition instead of their private payoff, where the welfare of the coalition is simply the sum of all its members' payoffs. On the contrary, a player outside the coalition, commonly called a free–rider, always maximises its own payoff. The difference between coalition members and non–members, therefore, translates into a difference in their respective optimization objectives:

$$\begin{aligned}
&\max_{q_i} \pi_i(Q, q_i), \quad \forall i \in N \setminus C, C \subseteq N; \\
&\max_{q_i} \sum_{j \in C} \pi_j(Q, q_j), \quad \forall i \in C, C \subseteq N;
\end{aligned}$$

with the former equation defining the objective function of a free–rider and the latter the one of a coalition member. Given a coalition C, we can insert equation (1) into the objective functions of members and non–members of C and solve for the optimal abatement of each player. This simply requires to take the derivative of the objective function of each player for her own level of abatement, q_i, equate it to zero and solve for q_i:

$$q_i^* = \frac{b_i - s_i \sum\limits_{j \in N \setminus i} q_j}{s_i + c_i}, \quad \forall i \in N \setminus C, \tag{2}$$

$$q_i^* = \frac{\sum\limits_{j \in C} b_j - \sum\limits_{j \in C} s_j \sum\limits_{j \in N \setminus i} q_j}{\sum\limits_{j \in C} s_j + c_i}, \quad \forall i \in C. \tag{3}$$

[6]The possibility of multiple coalitions would make the game a partition game.

Equations (2) and (3) define the optimal response function of each player to the abatement of the other players given coalition C. By forming a system of equations with the optimal response function of each player in N, an analytic solution for each q_i can be found. This is the optimal equilibrium level of abatement of each player in the game, from which global abatement Q can be easily retrieved. In Section A1 in the Appendix, the analytic solutions for the non–cooperative case (when no coalition forms) and for the grand coalition, when $C = N$, are reported.

The same procedure must be carried on for each possible coalition C, namely for each subset of N. Coalitions are then evaluated in order to find the ones that are stable. Stability is an equilibrium concept first envisaged for cartels (d'Aspremont et al., 1983) that translates the Nash equilibrium into a coalitional setting. A coalition is said to be stable if it is both externally stable, meaning that no free–riders has an incentive to join the coalition, and internally stable, meaning that no coalition member has an incentive to leave the coalition. In mathematical terms:

$$\pi_i(C) \geq \pi_i(C \setminus i), \quad \forall i \in C. \qquad \text{Internal Stability}$$

$$\pi_i(C) \geq \pi_i(C \cup i), \quad \forall i \in N \setminus C. \qquad \text{External Stability}$$

Stability is generally very difficult to achieve in such type of games, particularly for large coalitions (Barrett, 1994; Carraro and Siniscalco, 1993; Diamantoudi and Sartzetakis, 2006). Internal stability is the condition that more often fails to be met. A way of lessening this condition is to consider in its place potential internal stability (PIS), according to which a coalition is stable if it generates enough wealth so that each of its members can be made at least as well off as in case of acting as free–rider. Basically, a PIS coalition can be made internally stable through transfers. Its mathematical definition is the following:

$$\sum_{i \in C} \pi_i(C) \geq \sum_{i \in C} \pi_i(C \setminus i), \quad \forall i \in C. \qquad \text{PIS}$$

Stable coalitions, therefore, are the ones having the possibility to form and to remain in place, namely, not to dissolve. Similarly, PIS and externally stable coalitions can achieve the same results if opportune transfers are implemented. The other coalitions, instead, are bounded to dissolve either for the entrance of new members or, more likely, for the exit of unsatisfied members. Stable or PIS stable coalitions may then be ranked on the base of the total amount of public good provided.

3.2 The Barrett Model with Proportional Contributions

Reciprocity is that strategy for which a player conditions her cooperative behaviour on observing cooperation from her counterpart. Willing to translate this concept into a mechanism applicable to a public good game, we propose the idea that cooperating players abate a predefined proportion of average abatement (from which the name proportional contributions). This way, the more is the abatement of other players, the higher will be the abatement of player i, and vice–versa, insofar resembling the idea of reciprocity. The payoff of a player accepting the proportional contributions mechanism will be:

$$\pi_i = b_i Q - \frac{s_i}{2} Q^2 - \frac{c_i}{2} \left(\rho_i \sum_{j \in R} \frac{q_j}{|R|} \right)^2, \quad \forall i \in R; \tag{4}$$

where $\rho_i \in [0, |R|]$ is the parameter stating the proportion of abatement that player i must provide. Such value is multiplied for the average abatement: $\sum_{j \in R} \frac{q_j}{|R|}$, including the abatement of player i herself. Clearly, for $\rho_i < 1$, player i will provide a share of abatement below average. A natural condition to add is $\sum_{j \in R} \rho_j = |R|$, namely that the sum of $\boldsymbol{\rho}$ elements should be equal to the cardinality of R. This implies that, on average, each player's contribution to abatement will be equal to the average abatement: $\frac{1}{|R|} \sum_{i \in R} \rho_i \sum_{j \in R} \frac{q_j}{|R|} \equiv \sum_{j \in R} \frac{q_j}{|R|}$ since $\sum_{i \in R} \rho_i$ simplifies to $|R|$. The payoff and behaviour of free–riders, in this case players not abiding to the proportional contribution rule, are identical to the standard case.

The introduction of proportional contributions brings significant changes to the standard model. First of all, it has to be noted the notational difference between C and R. The latter can also be considered as a coalition, but while in the standard case coalition members differ from outsiders in their objective function, here the difference stems from abiding to the proportional rule. Members of R will still maximize own payoff, but their final contribution of abatement, q_i, will be determined by the average contribution rather than by their own pledge: $q_i = \rho_i \sum_{j\in R} \frac{q_j}{|R|}, \forall i \in R$. Further note that global abatement is given by the sum of abatement of outsiders and members of R: $Q = Q_R + Q_F$, where $Q_R = \sum_{i\in R} \rho_i \sum_{j\in R} \frac{q_j}{|R|} = |R|\bar{q}_i$ and $\bar{q}_i = \sum_{j\in R} \frac{q_j}{|R|}$. With Q_R being the total contribution of players abiding to the proportional contributions mechanism, Q_F is the total contribution of free–riders: $Q_F = \sum_{i\in N\setminus R} q_i$.

For solving the game, the same logic applied in the standard case holds in the present one. Once setting $r = |R|$ for notational convenience, we can write the optimal response function of a player abiding to the proportional contribution rule:

$$q_i^* = \frac{rb_i \sum\limits_{j\in N} \rho_j}{s_i \left(\sum\limits_{j\in N} \rho_j\right)^2 + c_i\rho_i^2} - \sum_{j\in N\setminus i} q_j = \frac{r^2 b_i}{s_i r^2 + c_i \rho_i^2} - \sum_{j\in N\setminus i} q_j, \quad \forall i \in R. \tag{5}$$

However, compared to the standard case, we have the elements of the $\boldsymbol{\rho}$ vector to determine as well. In fact, setting such values exogenously, may lead to sub–optimality. This leads to a problem since (5) provides an equation for each player in R, but we have two unknowns for each player, namely q_i and ρ_i, therefore the system is under–determined. By bringing the q_i from the LHS of equation (5) to the right side, we can rewrite (5) as:

$$Q = \frac{r^2 b_i}{s_i r^2 + c_i\rho_i^2} \to Q_R = \frac{r^2 b_i}{s_i r^2 + c_i \rho_i^2} - Q_F, \quad \forall i \in R.$$

The global abatement of free–riders, Q_F, is given by the sum of abatement of all free–riders, then it is completely determined by their optimal response functions, namely by (2). The other component of Q, as to say the global abatement provided by the players participating to the proportional contributions mechanism, Q_R, can be treated as a single variable rather than as the sum of the pledges of players in R. Therefore, we have $r + 1$

unknowns, one $\boldsymbol{\rho}$ element for each player in R plus Q_R. From (5), we have r equations, but adding the condition $\sum_{i \in R} \rho_i = r$ provides the additional equation to render the system squared. Therefore, the final system of equations to find all required unknowns is:

$$
\begin{aligned}
&Q_R = \frac{r^2 b_i}{s_i r^2 + c_i \rho_i^2} - Q_F, \quad \forall i \in R, \\
&q_i = \frac{b_i - s_i \left(Q_R + \sum\limits_{j \in N \setminus R \cup i} q_j \right)}{s_i + c_i}, \quad \forall i \in N \setminus R, \\
&\sum_{i \in R} \rho_i = r.
\end{aligned}
$$

Note that the single contribution to abatement of a player in R can be retrieved, once solved this system of equation, as $q_i^* = \rho_i^* \frac{Q_R^*}{r}, \forall i \in R$. The analytic solution in case of two players, both abiding to the proportional contributions mechanism, is present in Section A2 in the Appendix.

Despite R is not a proper coalition according to the standard definition since its members are still maximizing the private rather than the collective payoff, abiding to the proportional contributions mechanism still requires the willingness to give away the possibility of independently choosing the level of abatement. Therefore, R can still be analysed in terms of stability as if it was a proper coalition. Further note that, in case of identical players, when $\rho_i = 1, \forall i \in R$, the optimal level of abatement of each player is identical in the Barrett model with the grand coalition forming and under full participation to the proportional contributions mechanism[7].

3.3 The Proportional Contributions Mechanism with F&S Preferences

Reciprocity may be considered a rational strategy, and the results of the TIT–FOR–TAT strategy in the Axelrod tournament are a clear proof (Axelrod, 1980), but it may also be considered as a strategy motivated by fairness considerations. By reciprocating, a player rewards cooperative actions and sanctions non–cooperative ones. It is interesting, therefore,

[7] In particular $q_i^* = \frac{n b_i}{c_i + n^2 s_i}, \forall i \in C \quad \text{and} \quad \forall i \in R$.

to consider the proportional contributions mechanism together with preferences that include a fairness component. Following Vogt (2016) and Rogna and Vogt (2022), we use Fehr and Schmidt (1999) preferences due to their popularity and ability in explaining deviations to standard preferences observed in experimental settings (Fehr and Schmidt, 2006).

The F&S utility function subtracts to the material payoff of a player the cost of advantageous and disadvantageous inequality, under the assumption that the player is inequality averse. Following is the mathematical definition of the F&S utility function for a generic player i in N:

$$U_i = \pi_i - \frac{\alpha_i}{n-1} \sum_{m \in I^+} (\pi_m - \pi_i) - \frac{\beta_i}{n-1} \sum_{r \in I^-} (\pi_i - \pi_r), \quad \forall i \in N. \tag{6}$$

In equation (6), α and β are the two key parameters representing, respectively, the cost of disadvantageous and of advantageous inequality aversion. The sets I^+ and I^- include the players with, respectively, a payoff higher and lower than i: $j \in I^+ \Rightarrow \pi_j > \pi_i$ and $j \in I^- \Rightarrow \pi_j < \pi_i$. Given this, it is easy to see that both the summations in (6) are positive. The higher the difference between the payoff of i and the ones of players in I^+, the higher the cost for i in terms of disadvantageous inequality aversion. *Muta mutandis*, the same holds for advantageous inequality aversion (the β term). This implies that a more egalitarian distribution of final wealth increases the utility of player i. By assumption, largely confirmed by experimental evidence (Fehr and Schmidt, 1999), we have $\alpha_i \geq \beta_i$.

By inserting the definition of the payoff of a generic player i as given in equation (1) into (6), we can recreate the Barrett model but under F&S preferences. If, instead, we use equation (1) for defining the payoff of free–riders and (4) to define the payoff of players in R, we obtain a model with the proportional contributions mechanism and F&S preferences. Their solution is analogous to the one for their counterparts without F&S preferences, despite the computations being more cumbersome. In Section A3, in the Appendix, the analytic solutions for the case of F&S preferences with the proportional contributions mechanism are shown.

4 Simulations' Results

Comparing analytically the standard case, the introduction of the proportional abatement mechanism and this last case with F&S preferences is almost impossible due to the complexity of the results and to the number of parameters involved. Even numerical simulations, when parameter values are chosen arbitrarily, are problematic. Again, the main concern rests on the relatively high number of parameters. The result with a single set of parameters, in fact, is of scarce interest since it cannot be generalized. Parameter values should then be systematically varied in order to obtain generalizable results with the aim of finding parameter spaces where one case clearly prevails over the others. A high number of parameters, however, requires to perform such systematic variation in several dimensions, rendering very difficult the analysis of results. A third option, therefore, is to try to obtain numerical parameter values that are "realistic". In such occurrence, even results obtained with a specific set of parameters have useful implications since this should be the set of values of a real case scenario.

We try to follow this last option for our simulations by deriving the parameter values of the Barrett model from the popular RICE model of Nordhaus and Yang (1996) in its 2013 version (Nordhaus and Sztorc, 2013). The RICE v2013 model divides the world into twelve groups of countries/regions. For our simulations, however, due to complexity of the equations in case of F&S preferences with the proportional contributions mechanism, we have aggregated the world into four regions: High Income Countries (HIC), Mid Income Countries (MID), Developing Countries (DC) and China (CHN). Following is a brief explanation of the method used to derive the parameter values of our model.

4.1 Simulations' Settings

Once performed the aggregation into four regions, we have run the RICE model for a period of 100 years (from 2015 to 2215 with 10 years of time step) fixing both the level of investments and the proportion of abatement of all players. The former has been fixed in all model runs at the same level, as to say the one obtained in the standard non–cooperative case. The level of proportional abatement, instead, has been varied in the range (0, 1) with a 0.1 step for each single region separately. For each model run we have recorded the present value of per–capita consumption (PCI) – net of environmental damages and abatement costs – and the expenditure in abatement of each region (ω). Finally, we have estimated the following simple OLS model without intercept:

$$PCI_{i,j} = b_i \Omega_j + s_i \frac{\Omega_j^2}{2} + c_i \frac{\omega_{i,j}^2}{2} + u_{i,j}, \forall i \in \{\text{HIC, MIC, DC, CHN}\}\,; \tag{7}$$

where $\Omega_j = \sum_{i \in N} \omega_{i,j}$ is the sum of all abatement expenditures of all regions. Note that j indicates the values of a single model run. By running a regression for each region, it is possible to estimate b, s and c, the parameters of the Barrett model, for each player. Note that PCI and ω, as defined in equation (7), should represent faithfully their counterparts in the Barrett model, namely π and q. Table 1 reports the estimated coefficients for each region, all significant at the standard 5% level. Once retrieved the parameter values for our

Table 1: Estimated Parameter Values for the Barrett Model

	b	s	c
HIC	0.21302	-0.03387	-0.0047997
MIC	0.18480	-0.02972	-0.0037918
DC	0.04420	-0.00716	-0.0006068
CHN	0.10552	-0.01675	-0.0024523

model, the simulation and the benchmarking of cases can be performed.

4.2 Results and Discussion

A coalitional game with four players gives rise to 16 non–empty coalitions of which four are singletons, representing therefore the non–cooperative case. In presence of the proportional contributions mechanism, we have the same structure although with improper coalitions. As mentioned, we are particularly interested in finding coalitions that are stable (both internally and externally) or PIS stable.

Table 2 summarizes our findings for the case of standard preferences, showing the number of stable coalitions by stability criteria and by coalition size. Furthermore, it displays the maximum amount of abatement reached by stable coalitions. Such amount represents the expenditures for abatement undertaken by all regions (including free–riders) under a given coalition and it is expressed in trillions of dollars. Being maximal, such amount, when more than one stable coalition exists under a given criteria, refers to the coalition with the highest global expenditures for the public good.

Table 2: Number of Stable Coalitions by Stability Criteria and Coalition Size with Achieved Maximum Abatement

	Coalition Size			
	2 Players	**3 Players**	**Grand Coalition**	**Max. Global Q (T$)**
		Standard Case		
Stable	0	0	0	/
Internally Stable	0	0	0	/
Externally Stable	3	3	1	6.23
PIS	1	0	0	6.199
		Proportional Contributions Mechanism		
Stable	0	0	0	/
Internally Stable	1	0	0	5.86
Externally Stable	4	4	1	6.172
PIS	1	2	0	6.157

Under standard preferences, it is possible to observe that no fully stable coalitions are possible, neither in the standard case nor with the proportional contributions mechanism. The lack of internal stability is the principal obstacle. In fact, no internally stable coalitions are present in the standard case. The proportional contributions mechanism allows for an internally two players coalition between China and Developing Countries, that, however, fails to be externally stable. Even when considering external stability, the proportional contributions mechanism performs slightly better, satisfying this condition for 4 coalitions of 2 players out of six (they are 3 in the standard case) and for all four coalitions with 3 players (3 under standard preferences). The grand coalition always meet this condition, therefore it is externally stable in all considered cases. The largest difference between the standard case and the introduction of the alternative mechanism happens when PIS coalitions are considered. In both cases we have a single 2 players coalition being stable, but the proportional contributions mechanism further stabilizes 2 coalitions with size of three players, namely (HIC, DC, CHN) and (MIC, DC, CHN).

The larger stability obtained with the proportional contributions mechanism comes at a cost. Specifically, the contribution to the production of the public good abatement is substantially lower with this mechanism that under the standard optimization of the coalition welfare. If we look at the external stability rows in Table 2, the level of maximum global Q refers to the grand coalition, that clearly grants the highest contribution by preventing free–riding. The difference in the values related to the standard case and to the proportional contributions mechanism is significant: 6.230 versus 6.172. Moreover, we can see that the value of Q of the grand coalition in case of proportional contributions is lower than what achieved in the standard case by a 2 players PIS coalition: $6.199 > 6.172$. The coalition of HIC and DC, that is PIS stable in the standard case, reaches a higher level of public good production compared to the coalition of all regions with proportional contributions.

When F&S preferences are introduced, there are two more parameters to consider, namely the inequality aversion parameters α and β. Fehr and Schmidt (1999) provide a range of values as retrieved from various experiments: $\alpha \in [0, 4.5]$ and $\beta \in [0, 0.6]$. Using median or average values is a possibility, but the results will be confined to this specific combination, that may not be more realistic than other possible combinations. Our approach has been to systematically vary the values of such parameters inside the given plausible ranges with the aim of finding sub–spaces of values with similar patterns. This renders the presentation of results more difficult, since we have tens of simulations cumbersome to summarize in a table. For this reason, and given that under F&S preferences it is actually possible to find PIS stable grand coalitions, only the stability of the grand coalition will be discussed here.

In general, the role of β is far less important than the one of α to determine the PIS stability of the grand coalition (Rogna and Vogt, 2020, 2022). Therefore, we can fix it to its median value, equal to 0.288, since other values in the [0, 0.6] range do not cause dramatic changes. Full stability is never reached both in the simple case of F&S preferences and when the proportional contributions mechanism is added. However, PIS stability is obtained in both cases for certain ranges of α values. A distinction has to be made. In fact, in several of our simulations, we have obtained negative qs, namely negative abatement for certain regions. Although they may seem unrealistic, if we think abatement commitments as relative to a certain given level of pollution, a negative abatement would then imply to increase that level of pollution. In the Kyoto protocol, for example, the countries targets are set in terms of percentage reduction of pollution compared to the 1990 level. Therefore, a negative abatement would be the increase of pollution of a country compared to its 1990 level.

Table 3: The Space of α Values Allowing for PIS Stable Grand Coalitions with F&S Preferences

Cases	**α space with PIS stable grand coalitions**
	Only positive qs
F&S Preferences	–
F&S Preferences with proportional contributions	$1.40 \leq \alpha \leq 1.46, 2.90 \leq \alpha \leq 4.00$
	Including negative qs
F&S Preferences	$0.833, 1.70 \leq \alpha \leq 2.30, 2.75 \leq \alpha \leq 4.00$
F&S Preferences with proportional contributions	$1.40 \leq \alpha \leq 2.00, 2.60 \leq \alpha \leq 4.00$

Table 3 resumes our findings. In particular, we can see that there are no PIS stable grand coalition under simple F&S preferences if we admit only positive qs. By adding proportional contributions, instead, we have PIS stable grand coalitions for $1.40 \leq \alpha \leq 1.46$

and for $2.90 \leq \alpha \leq 4.00$. When we further consider negative qs, F&S preferences alone are sufficient to sustain PIS grand coalitions. This happens for α equal to 0.833, the median value in Fehr and Schmidt (1999) and for $\alpha \in [1.7, 2.3]$ and $\alpha \in [2.75, 4]$. The introduction of proportional contributions still increases the possibility of PIS stable grand coalitions. In fact, the α spaces allowing for PIS stability are enlarged: $\alpha \in [1.4, 2]$ and $\alpha \in [2.6, 4]$. Only for $\alpha = 0.833$ and for $\alpha \in (2, 2.3]$ the absence of the proportional contributions mechanism is beneficial for PIS stability. On the contrary, for $\alpha \in [1.4, 1.7]$ and $\alpha \in [2.6, 2.9]$, it is the proportional contribution mechanism that allows for PIS stability while F&S preferences alone fail to do so.

Table 4: Global abatement (Q) under F&S Preferences with and without Proportional Contributions

	With Reciprocity	**Without Reciprocity**		**With Reciprocity**	**Without Reciprocity**
$\alpha = 1.7$	5.980	6.006	$\alpha = 3.2$	5.369	6.176
$\alpha = 1.8$	5.917	5.978	$\alpha = 3.3$	5.297	6.125
$\alpha = 1.9$	5.820	5.940	$\alpha = 3.4$	5.424	6.077
$\alpha = 2.0$	5.695	5.899	$\alpha = 3.5$	5.523	6.029
$\alpha = 2.1$	5.510	5.830	$\alpha = 3.6$	5.602	5.979
$\alpha = 2.8$	3.620	6.556	$\alpha = 3.7$	5.666	5.926
$\alpha = 2.9$	5.001	6.401	$\alpha = 3.8$	5.719	5.867
$\alpha = 3.0$	4.938	6.305	$\alpha = 3.9$	5.762	5.801
$\alpha = 3.1$	5.188	6.234			

It has to be noted that, also under F&S preferences, the level of global abatement Q is sensibly reduced when proportional contributions are allowed. This is clearly shown in Table 4, where the values of Qs in presence of the grand coalition is reported for different α values. In all cases, the reciprocity mechanism leads to a lower level of global abatement compared to F&S preferences alone. The difference seems to be less marked than the standard preference case, but this is probably due to the fact that, for most values of α, F&S preferences already lead to a lower level of Q compared to standard preferences. Therefore, also in case of F&S

preferences, it is possible to conclude that proportional contributions increase cooperation at the cost of a lower level of the public good.

5 Conclusions

The cooperation for producing a public good, such as the reduction of global pollution, is difficult to obtain. The several failures of climate negotiations are a clear example that has been largely anticipated by the game theoretical analysis of these types of agreement. At the same time, there are examples of successful agreements, particularly in the field of international trade, that are difficult to explain if standard economic theory is adopted. One possibility to overcome these theoretical shortcomings is to adopt other–regarding preferences in public good games with a coalitional structure. This may help in better representing the motivations of actors involved in international agreements.

However, when a popular representation of other–regarding behaviour, namely F&S preferences, is introduced in a coalitional game of public good production, the cooperation enhancing effect is rather modest. Transfers may further help to increase the space of cooperation, but the grand coalition generally remains unstable. The present paper investigates another possibility, namely the introduction of a mechanism that is rooted into the principle of reciprocity: proportional contributions. A player abiding to this mechanism will provide an amount of the public good that is proportional to the average contribution proposed by all actors abiding to the mechanism. Therefore, its contribution will be increased by an increase in contributions by the other players and vice–versa.

The proportional contributions mechanism is compared to the standard formulation of a coalitional game for the production of a public good. In particular, we adopt the popular abatement game of Barrett parameterized on the data of the RICE model. The mechanism is benchmarked both under the assumption of standard preferences and under the introduction of F&S preferences. In both cases there is a significant increase in cooperation. Under standard preference, however, it is still not possible to stabilize the grand coalition, not even when allowing for transfers (PIS stability). The introduction of F&S preferences, instead, allows for PIS stable coalitions even without the proportional contributions mechanism in place. However, the introduction of this last enlarges the space of α values, with α being the degree of disadvantageous inequality aversion, for which cooperation is possible. In all cases, however, the introduction of the mechanism implies a sensible reduction of global abatement. Therefore, the increase in cooperation is reached by sacrificing a portion of public good.

References

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(1):3–25.

Bagwell, K. and Staiger, R. W. (1999). An economic theory of GATT. *American Economic Review*, 89(1):215–248.

Bagwell, K. and Staiger, R. W. (2004). Multilateral trade negotiations, bilateral opportunism and the rules of GATT/WTO. *Journal of International Economics*, 63(1):1–29.

Bagwell, K. and Staiger, R. W. (2016). The design of trade agreements. In *Handbook of Commercial Policy*, volume 1, pages 435–529. Elsevier.

Bakalova, I. and Eyckmans, J. (2019). Simulating the impact of heterogeneity on stability and effectiveness of international environmental agreements. *European Journal of Operational Research*, 277(3):1151–1162.

Barrett, S. (1994). Self-enforcing International Environmental Agreements. *Oxford Economic Papers*, 46:878–894.

Bolton, G. E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193.

Bolton, G. E. and Ockenfels, A. (2005). A stress test of fairness measures in models of social utility. *Economic Theory*, 25(4):957–982.

Breton, M. and Garrab, S. (2014). Evolutionary farsightedness in international environmental agreements. *International Transactions in Operational Research*, 21(1):21–39.

Carraro, C. and Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, 52(3):309–328.

Chander, P. and Tulkens, H. (1995). A Core-theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution. *International Tax and Public Finance*, 2(2):279–293.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J., and Weymark, J. A. (1983). On the stability of collusive price leadership. *Canadian Journal of Eonomics*, 16(1):17–25.

Diamantoudi, E. and Sartzetakis, E. S. (2006). Stable international environmental agreements: An analytical approach. *Journal of Public Economic Theory*, 8(2):247–263.

Diamantoudi, E. and Sartzetakis, E. S. (2015). International environmental agreements: coordinated action under foresight. *Economic Theory*, 59(3):527–546.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Fehr, E. and Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the Economics of Giving, Altruism and Reciprocity*, 1:615–691.

Fischbacher, U. and Gachter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, 100(1):541–56.

Freund, C. L. (2003). Reciprocity in free trade agreements. *Available at SSRN 636423.*

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, 3(4):367–388.

Isaac, R. M., Walker, J. M., and Williams, A. W. (1994). Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups. *Journal of Public Economics*, 54(1):1–36.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of business*, pages S285–S300.

Keohane, R. O. (1982). The demand for international regimes. *International Organization*, 36(2):325–355.

Lamp, N. (2017). The 'development' discourse in multilateral trade lawmaking. *World Trade Review*, 16(3):475–500.

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

Lange, A. and Vogt, C. (2003). Cooperation in international environmental negotiations due to a preference for equity. *Journal of Public Economics*, 87(9):2049–2067.

McGinty, M. (2007). International environmental agreements among asymmetric nations. *Oxford Economic Papers*, 59(1):45–62.

Nordhaus, W. and Sztorc, P. (2013). Dice 2013r: Introduction and user's manual. *Yale University and the National Bureau of Economic Research, USA*.

Nordhaus, W. D. and Yang, Z. (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies. *The American Economic Review*, 86(4):741–765.

Nyborg, K. (2018). Social norms and the environment. *Annual Review of Resource Economics*, 10:405–423.

Pavlova, Y. and De Zeeuw, A. (2013). Asymmetries in international environmental agreements. *Environment and Development Economics*, 18(1):51–68.

Rhodes, C. (1989). Reciprocity in trade: The utility of a bargaining strategy. *International Organization*, 43(2):273–299.

Rogna, M. and Vogt, C. (2020). *Coalition formation with optimal transfers when players are heterogeneous and inequality averse*. Number 865. Ruhr Economic Papers.

Rogna, M. and Vogt, C. (2022). Optimal climate policies under fairness preferences. *Climatic Change*, 174:1–25.

van der Pol, T., Weikard, H.-P., and van Ierland, E. (2012). Can altruism stabilise international climate agreements? *Ecological Economics*, 81:112–120.

Vogt, C. (2016). Climate coalition formation when players are heterogeneous and inequality averse. *Environmental and Resource Economics*, 65(1):33–59.

Weikard, H.-P., Wangler, L., and Freytag, A. (2015). Minimum participation rules with heterogeneous countries. *Environmental and Resource Economics*, 62(4):711–727.

Appendix

A1 Analytic Solutions for the Optimal Level of Abatement of Coalition Members and Non–Members in the Standard Barrett Model (No– and Full–Cooperation)

Non–cooperative solution:

$$q_i^* = \frac{b_i\left(\prod\limits_{j\in N\setminus i} c_j + \sum\limits_{j\in N\setminus i}\prod\limits_{k\in N\setminus i,j} c_k s_j\right) - s_i \sum\limits_{j\in N\setminus i}\prod\limits_{k\in N\setminus i,j} c_k b_j}{\prod\limits_{j\in N} c_j + \sum\limits_{j\in N}\prod\limits_{k\in N\setminus j} c_k s_j}, \quad \forall i \in N. \tag{A1}$$

Cooperative solution:

$$q_i^* = \frac{\sum\limits_{j\in N} b_j \prod\limits_{k\in N\setminus i} c_k}{\prod\limits_{j\in N} c_j + \sum\limits_{j\in N}\sum\limits_{k\in N}\prod\limits_{h\in N\setminus k} c_h s_j}, \quad \forall i \in N. \tag{A2}$$

A2 Analytic Solution for the Proportional Contributions Mechanism in Case of Two Players both Belonging to R

$$\rho_1 = n\frac{b_1c_2 \pm \sqrt{-b_1^2c_2s_2 + b_1b_2c_1c_2 + b_1b_2c_1s_2 + b_1b_2c_2s_1 - b_2^2c_1s_1}}{b_1c_2 - b_2c_1},$$
$$\rho_2 = n\frac{b_2c_1 \pm \sqrt{-b_1^2c_2s_2 + b_1b_2c_1c_2 + b_1b_2c_1s_2 + b_1b_2c_2s_1 - b_2^2c_1s_1}}{b_2c_1 - b_1c_2}.$$

Note that $Q = Q_R$ since $N \setminus R = \emptyset \Rightarrow Q_F = 0$:

$$Q = \frac{b_1c_1c_2^2 - b_1c_1c_2s_2 + b_1c_2^2s_1 + b_2c_1^2c_2 + b_2c_1^2s_2 - b_2c_1c_2s_1 \pm 2c_1c_2\sqrt{-b_1^2c_2s_2 + b_1b_2c_1c_2 + b_1b_2c_1s_2 + b_1b_2c_2s_1 - b_2^2c_1s_1}}{c_1^2c_2^2 + 2c_1^2c_2s_2 + c_1^2s_2^2 + 2c_1c_2^2s_1 - 2c_1c_2s_1s_2 + c_2^2s_1^2}.$$

A3 Analytic Solutions with Fehr and Schmidt Preferences and Proportional Contributions Mechanism

By inserting the Barrett payoff function expressed in equation (1) into the F&S utility function as in equation (6), we have the explicit equation of the utility of a generic player i:

$$
\begin{aligned}
&U_i = b_i \sum_{k\in N} \frac{\bar{q}}{n}\rho_k - \frac{s_i}{2}\left(\sum_{k\in N}\frac{\bar{q}}{n}\rho_k\right)^2 - \frac{c_i}{2}\left(\rho_i\bar{q}\right)^2 \\
&-\frac{\alpha_i}{n-1}\sum_{j\in I^+}\left(b_j \sum_{k\in N}\frac{\bar{q}}{n}\rho_k - \frac{s_j}{2}\left(\sum_{k\in N}\frac{\bar{q}}{n}\rho_k\right)^2 - \frac{c_j}{2}\left(\rho_j\bar{q}\right)^2 - \left(b_i\sum_{k\in N}\frac{\bar{q}}{n}\rho_k - \frac{s_i}{2}\left(\sum_{k\in N}\frac{\bar{q}}{n}\rho_k\right)^2 - \frac{c_i}{2}\left(\rho_i\bar{q}\right)^2\right)\right) \\
&-\frac{\beta_i}{n-1}\sum_{j\in I^-}\left(b_i \sum_{k\in N}\frac{\bar{q}}{n}\rho_k - \frac{s_i}{2}\left(\sum_{k\in N}\frac{\bar{q}}{n}\rho_k\right)^2 - \frac{c_i}{2}\left(\rho_i\bar{q}\right)^2 - \left(b_j\sum_{k\in N}\frac{\bar{q}}{n}\rho_k - \frac{s_j}{2}\left(\sum_{k\in N}\frac{\bar{q}}{n}\rho_k\right)^2 - \frac{c_j}{2}\left(\rho_j\bar{q}\right)^2\right)\right).
\end{aligned}
$$

With this being the explicit utility function of a generic player i, we can find the optimal level of q_i by taking the first derivative of this function for q_i and equating it to zero for the case of a player not belonging to any coalitions and repeating the same procedure but over the sum of utilities of all players inside a coalition for a member of it. Following are reported the optimal response functions of a generic player i for the non–cooperative case and for the case of full cooperation:

$$
q_i = \frac{b_i - \frac{\beta_i}{n-1}\left(|I^-|b_k - \sum_{j\in I^-} b_j\right) - \frac{\alpha_i}{n-1}\left(\sum_{j\in I^+} b_j - |I^+|b_k\right)}{\sum_{k\in N} s_k + \frac{c_k\rho_k^2}{n^2} + \sum_{k\in N}\frac{\beta_k}{n-1}\left(\sum_{j\in K^-} s_j + \frac{c_j\rho_j^2}{n^2} - |K^-|s_k + \frac{c_k\rho_k^2}{n^2}\right) + \sum_{k\in N}\frac{\alpha_k}{n-1}\left(|K^+|s_k + \frac{c_k\rho_k^2}{n^2} - \sum_{j\in K^+} s_j + \frac{c_j\rho_j^2}{n^2}\right)} - \sum_{j\in N\setminus i} q_j, \forall i \in N.
$$

$$
q_i = \frac{b_i - \sum_{k\in N}\frac{\beta_k}{n-1}\left(|K^-|b_k - \sum_{j\in K^-} b_j\right) - \sum_{k\in N}\frac{\alpha_k}{n-1}\left(\sum_{j\in K^+} b_j - |K^+|b_k\right)}{\sum_{k\in N} s_k + \frac{c_k\rho_k^2}{n^2} + \sum_{k\in N}\frac{\beta_k}{n-1}\left(\sum_{j\in K^-} s_j + \frac{c_j\rho_j^2}{n^2} - |K^-|s_k + \frac{c_k\rho_k^2}{n^2}\right) + \sum_{k\in N}\frac{\alpha_k}{n-1}\left(|K^+|s_k + \frac{c_k\rho_k^2}{n^2} - \sum_{j\in K^+} s_j + \frac{c_j\rho_j^2}{n^2}\right)} - \sum_{j\in N\setminus i} q_j, \forall i \in C.
$$